

GenCore version: 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2004, 03:25:04 ; Search time 925 Seconds  
(without alignments)  
11198.630 Million cell updates/sec

Title: US-09-765-272A-65

Perfect score: 2290

Sequence: 1 TTGTCCTATGAACTGGTC.....TACTTGAAAAATCCAGGAT 2290

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARY

Result No.	Score	Query Match	Length	ID	Description
1	2285	99.8	2290	9	US-09-765-272-65
2	2244.8	98.0	2481	10	US-09-765-787-206
3	2220.2	97.0	2531	13	US-10-412-862-11
4	2220.2	97.0	2531	16	US-10-412-862-11
5	2220.2	97.0	2531	16	US-10-412-862-11
6	2220.2	97.0	2531	16	US-10-412-862-11
7	1999.6	87.3	2531	13	US-10-412-862-5
8	1999.6	87.3	2531	16	US-10-412-862-5
9	1999.6	87.3	2531	16	US-10-412-862-5
10	1999.2	87.3	1815	13	US-10-158-844-94
11	1856.8	81.1	2523	10	US-09-884-465A-3
12	1856.8	81.1	2647	10	US-09-884-465A-4
13	1376.6	60.1	2359	13	US-10-158-844-243
14	990.6	43.3	2451	13	US-10-412-862-9

15	990.6	43.3	2451	16	US-10-412-862-9	Sequence 9, Appli
16	990.6	43.3	2451	16	US-10-387-783-9	Sequence 9, Appli
17	987.6	43.1	2389	9	US-09-765-272-55	Sequence 55, Appli
18	794.8	34.7	973	13	US-10-158-844-355	Sequence 355, App
19	677	29.6	1684	13	US-10-158-844-258	Sequence 258, App
20	401	17.5	1342	9	US-09-765-272-181	Sequence 181, App
21	401	17.5	1455	10	US-09-769-787-246	Sequence 246, App
22	401	17.5	1455	10	US-09-769-787-246	Sequence 246, App
23	401	17.5	1455	13	US-10-412-862-7	Sequence 7, Appli
24	401	17.5	1455	16	US-10-412-862-7	Sequence 7, Appli
25	401	17.5	1455	16	US-10-387-783-7	Sequence 7, Appli
26	401	17.5	3120	10	US-09-884-465A-1	Sequence 1, Appli
27	401	17.5	5048	10	US-09-884-465A-2	Sequence 2, Appli
28	401	17.5	6867	13	US-10-158-844-192	Sequence 192, App
29	370	16.2	819	10	US-09-884-465A-257	Sequence 257, App
30	370	16.2	819	16	US-10-324-143-10	Sequence 10, Appli
31	368.4	16.1	816	16	US-10-324-143-4	Sequence 4, Appli
32	368.4	16.1	816	16	US-10-324-143-5	Sequence 5, Appli
33	363.6	15.9	2721	16	US-10-324-143-6	Sequence 6, Appli
34	346.6	15.1	2528	10	US-09-884-465A-9	Sequence 9, Appli
35	272.4	11.9	841	9	US-09-452-599-34	Sequence 34, Appli
36	272.4	11.9	841	15	US-10-121-120-34	Sequence 34, Appli
37	225	9.8	1146	10	US-09-769-736-23	Sequence 23, Appli
38	225	9.8	2469	10	US-09-769-736-17	Sequence 17, Appli
39	223.4	9.8	5215	10	US-09-252-088-13	Sequence 13, Appli
40	223.4	9.8	5215	16	US-10-340-792-13	Sequence 13, Appli
41	117	5.1	1455	10	US-09-769-736-71	Sequence 71, Appli
42	58.2	2.5	31124	13	US-10-087-192-463	Sequence 463, App
43	54.8	2.4	45980	10	US-09-957-956-6	Sequence 6, Appli
44	54	2.4	2232	15	US-10-087-464-45	Sequence 45, Appli
45	52.6	2.3	635	13	US-10-027-632-269927	Sequence 269927,
46	52.6	2.3	635	16	US-10-027-632-269927	Sequence 269927,
47	52.2	2.2	96596	12	US-10-052-483-70	Sequence 70, Appli
48	51.2	2.2	305	9	US-09-864-761-19262	Sequence 19262, A
49	51.2	2.2	496	9	US-09-864-761-2534	Sequence 2534, Ap
50	50.6	2.2	5928	13	US-10-221-714A-492	Sequence 492, App
51	50.6	2.2	5928	15	US-10-311-455-2212	Sequence 2212, Ap
52	50	2.2	1010	15	US-10-029-386-24901	Sequence 24901, A
53	49.8	2.2	122859	13	US-10-087-193-37	Sequence 37, Appli
54	49.6	2.2	276	9	US-09-864-761-20595	Sequence 20595, A
55	49.4	2.2	678	13	US-10-027-632-141716	Sequence 141716,
56	49.4	2.2	678	16	US-10-027-632-141716	Sequence 141716,
57	49.2	2.1	438	9	US-09-864-761-4988	Sequence 4988, Ap
58	49.2	2.1	2499	9	US-09-842-552-92	Sequence 92, Appli
59	49.2	2.1	2550	15	US-10-032-585-6026	Sequence 6026, Ap
60	48.8	2.1	408	9	US-09-360-352-1221	Sequence 1221, Ap
61	48.8	2.1	524	13	US-10-027-632-50439	Sequence 50439, A
62	48.8	2.1	524	16	US-10-027-632-50439	Sequence 50439, A
63	48.4	2.1	37285	13	US-10-087-192-49	Sequence 49, Appli
64	48.4	2.1	374849	13	US-10-087-192-1627	Sequence 1627, Ap
65	48.2	2.1	2276	13	US-10-027-632-110545	Sequence 110545,
66	48.2	2.1	2276	16	US-10-027-632-110545	Sequence 110545,
67	48.2	2.1	14798	15	US-10-311-455-1006	Sequence 1006, Ap
68	48	2.1	3912	13	US-10-282-122A-9928	Sequence 9928, Ap
69	48	2.1	96597	12	US-10-052-482-103	Sequence 1023, A
70	47.8	2.1	315	9	US-09-864-761-21723	Sequence 21723, A
71	47.8	2.1	7200	13	US-10-282-122A-33992	Sequence 33992, A
72	47.8	2.1	18846	9	US-09-815-242-8898	Sequence 8898, Ap
73	47.8	2.1	31096	8	US-08-781-986A-59	Sequence 59, Appli
74	47.8	2.1	31096	13	US-10-329-624-59	Sequence 59, Appli
75	47.4	2.1	1671	13	US-10-412-698B-1191	Sequence 1191, Ap
76	47.4	2.1	1671	16	US-10-374-780A-834	Sequence 834, App
77	47.2	2.1	536	15	US-10-029-386-19974	Sequence 19974, A
78	47.2	2.1	536	15	US-10-029-386-19974	Sequence 6243, Ap
79	47	2.1	143899	10	US-09-972-546-15	Sequence 15, Appli
80	46.8	2.0	646	13	US-10-424-599-39641	Sequence 39641, A
81	46.6	2.0	1482	9	US-09-070-327A-842	Sequence 842, App
82	46.4	2.0	327	9	US-09-864-761-28059	Sequence 28059, A
83	46.4	2.0	439	9	US-09-960-352-1009	Sequence 1009, Ap
84	46.4	2.0	193853	13	US-10-087-192-1663	Sequence 1663, Ap
85	46.2	2.0	1282	14	US-10-002-344A-89	Sequence 89, Appli
86	46.2	2.0	2392	15	US-10-084-817-312	Sequence 312, App
87	46.2	2.0	5631	13	US-10-282-122A-37677	Sequence 37677, A

88 46 2.0 522 14 US-10-101-487-71  
C 89 46 2.0 530 14 US-10-101-487-73  
90 46 2.0 554 14 US-10-101-487-69  
91 46 2.0 554 14 US-10-101-487-106  
92 45.8 2.0 2916 13 US-10-282-122A-17169  
93 45.8 2.0 5394 15 US-10-311-455-981  
94 45.4 2.0 1848 13 US-10-027-632-98794  
95 45.4 2.0 1848 13 US-10-027-632-98795  
96 45.4 2.0 1848 16 US-10-027-632-98794  
97 45.4 2.0 1848 16 US-10-027-632-98795  
98 45.2 2.0 423 9 US-09-864-761-18355  
99 45.2 2.0 68233 16 US-10-031-650-31  
100 45.2 2.0 118502 13 US-10-087-192-397  
101 44.8 2.0 4316 9 US-09-880-107-3713  
102 44.8 2.0 14006 15 US-10-311-455-1931  
103 44.8 2.0 3673778 15 US-10-312-841-2  
104 44.8 2.0 1680 10 US-09-298-523B-73  
105 44.6 1.9 1680 13 US-09-748-875-73  
106 44.6 1.9 20142 13 US-10-282-122A-7627  
107 44.6 1.9 286 15 US-10-029-386-18943  
108 44.4 1.9 569 13 US-10-424-599-102311  
109 44.4 1.9 1047 13 US-10-029-386-5187  
110 44.4 1.9 1047 13 US-10-335-977-822  
111 44.4 1.9 1050 13 US-10-335-977-822  
112 44.4 1.9 451 9 US-09-960-352-10262  
113 44.2 1.9 462 9 US-09-864-761-3829  
114 44.2 1.9 804 13 US-10-027-632-29676  
115 44.2 1.9 804 16 US-10-027-632-29676  
116 44.2 1.9 6668 15 US-10-311-455-1670  
117 44.2 1.9 106664 15 US-10-175-523-97  
118 44.2 1.9 554 13 US-10-027-632-64906  
119 44 1.9 554 16 US-10-027-632-64906  
120 44 1.9 1410 9 US-09-815-242-7436  
121 44 1.9 1410 13 US-10-335-977-823  
122 44 1.9 1410 13 US-10-335-977-824  
123 44 1.9 1410 13 US-10-335-977-824  
124 44 1.9 2522 9 US-09-842-552-98  
125 44 1.9 1691139 15 US-10-067-514-1  
126 44 1.9 1691139 16 US-10-419-723-1  
127 44 1.9 3673778 15 US-10-312-841-1  
128 43.8 1.9 483 9 US-09-920-300A-439  
129 43.8 1.9 483 14 US-10-032-528-439  
130 43.8 1.9 483 15 US-10-099-926-439  
131 43.8 1.9 554 9 US-09-796-692-4144  
132 43.8 1.9 554 15 US-10-040-862-4144  
133 43.8 1.9 554 16 US-10-057-475B-4144  
134 43.8 1.9 554 16 US-10-154-884B-4144  
135 43.8 1.9 765 14 US-10-141-627-3  
136 43.8 1.9 855 14 US-10-141-627-3  
137 43.8 1.9 2520 17 US-10-389-566-205  
138 43.8 1.9 3061 9 US-09-833-790-414  
139 43.8 1.9 3061 9 US-09-880-107-3751  
140 43.8 1.9 3061 13 US-10-382-360-365  
141 43.8 1.9 3071 9 US-09-880-107-1651  
142 43.8 1.9 3253 10 US-09-814-353-20312  
143 43.8 1.9 8149 15 US-10-311-455-1123  
144 43.8 1.9 126413 13 US-10-087-192-1831  
145 43.6 1.9 5195 15 US-10-311-455-893  
146 43.6 1.9 3673778 15 US-10-312-841-1  
147 43.4 1.9 767 13 US-10-424-599-70401  
148 43.4 1.9 2014 9 US-09-842-552-22  
149 43.4 1.9 2065 15 US-10-106-698-1489  
150 43.4 1.9 3827 13 US-10-433-544-16

## ALIGNMENT:

RESULT 1  
US-09-765-272-65  
; Sequence 65, Application US/09765272  
; Patent No. US20020061545A1  
; GENERAL INFORMATION:

APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,272  
FILING DATE: 22-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/961,083  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2290 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-09-765-272-65  
Query Match 99.8%; Score 2285; DB 9; Length 2290;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTGTTCTATGAACCTTGGTCGTACCAAGCTGGTCAGGTTAAGAAAGAGTCTATCGAGT 60  
Db 1 TTGTTCTATGAACCTTGGTCGTACCAAGCTGGTCAGGTTAAGAAAGAGTCTATCGAGT 60  
QY 61 TTCTTATATAGATCGTATCAGGCTGCTCAAAAGCGAGAACTTGACACCAGATGAAGT 120  
Db 61 TTCTTATATAGATCGTATCAGGCTGCTCAAAAGCGAGAACTTGACACCAGATGAAGT 120  
QY 121 CAGTAAGAGGGAGGGGATCAACGCCGAACTTATCAAGATTACGGATCAAGGTTA 180  
Db 121 CAGTAAGAGGGAGGGGATCAACGCCGAACTTATCAAGATTACGGATCAAGGTTA 180  
QY 181 TGTGACCTCTCATGGAGACCACTTATCTATTAATGCGCAAGGTTCTTATGATGCCAT 240  
Db 181 TGTGACCTCTCATGGAGACCACTTATCTATTAATGCGCAAGGTTCTTATGATGCCAT 240  
QY 241 CATCAGTGAAGAGCTCTCTCATGAAGAGATCCGAATTTATCAGTTTGAAGGATTCAGACATGT 300  
Db 241 CATCAGTGAAGAGCTCTCTCATGAAGAGATCCGAATTTATCAGTTTGAAGGATTCAGACATGT 300  
QY 301 CAATGAATCAAGGGTGGTTATGTCTATTAAGGTTAAACCGTAAATATCTATGNTTACCTTAA 360  
Db 301 CAATGAATCAAGGGTGGTTATGTCTATTAAGGTTAAACCGTAAATATCTATGNTTACCTTAA 360  
QY 361 CGATGACGCTCATGGGATTAATATTCGACAAAGAGAGATTAAACGTCAGACGAGCA 420  
Db 361 CGATGACGCTCATGGGATTAATATTCGACAAAGAGAGATTAAACGTCAGACGAGCA 420  
QY 421 ACGCAGTCAATATCAATCACTCAAGGACAGATAATGCTGTTCTGTCAGCAGACGCCAAGG 480  
Db 421 ACGCAGTCAATATCAATCACTCAAGGACAGATAATGCTGTTCTGTCAGCAGACGCCAAGG 480

481 ACCTTATACACCGGATGATGGGTATATCTTCAATGTCATCTGATATCATGAGGACACGGG 540  
481 ACCTTATACACCGGATGATGGGTATATCTTCAATGTCATCTGATATCATGAGGACACGGG 540  
541 TGATGCTTATATGCTTCCCTCAGCGGACCTTTACCAATTTACATTTCTTAAAGATGAGTTATC 600  
541 TGATGCTTATATGCTTCCCTCAGCGGACCTTTACCAATTTACATTTCTTAAAGATGAGTTATC 600  
601 AGCTAGCGATGATGCTGCTGAGAGCCCTTTTGGAAATGGGAAGGAGGATCTCGTCCCTTC 660  
601 AGCTAGCGATGATGCTGCTGAGAGCCCTTTTGGAAATGGGAAGGAGGATCTCGTCCCTTC 660  
661 TTCAAGTCTTCTAGTTTAAATGCAAAATCCAGTCAACCAAGATGTCAGAGAAACCAAAATCT 720  
661 TTCAAGTCTTCTAGTTTAAATGCAAAATCCAGTCAACCAAGATGTCAGAGAAACCAAAATCT 720  
721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGA 780  
721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGA 780  
781 ATTGTATGCTAAACCTTATCAGAAAGCCCTTGTGGAAATCTGATGGCCTTATTTTGCACCC 840  
781 ATTGTATGCTAAACCTTATCAGAAAGCCCTTGTGGAAATCTGATGGCCTTATTTTGCACCC 840  
841 AGCGCAAAATCACAAGTTCGAACCGCCAGAGTGTAGCTGTCCCTCATGTGTAAACCATTTACCA 900  
841 AGCGCAAAATCACAAGTTCGAACCGCCAGAGTGTAGCTGTCCCTCATGTGTAAACCATTTACCA 900  
901 CTTTATCCCTTATGAAACAAATGTCTGAAATGGGAAACCAATTTGCTCGTATTTATCCCT 960  
901 CTTTATCCCTTATGAAACAAATGTCTGAAATGGGAAACCAATTTGCTCGTATTTATCCCT 960  
961 TCGTATGCTTCAACCAATGGGTACAGTTCAGACCAAGCAACCAAGTTCAGCAATTC 1020  
961 TCGTATGCTTCAACCAATGGGTACAGTTCAGACCAAGCAACCAAGTTCAGCAATTC 1020  
1021 GACTCCGGAACTTAGTCCCAAGTCCGCAACCTGACCAAAATCTTCAACCAAGTTCAGCAATTC 1080  
1021 GACTCCGGAACTTAGTCCCAAGTCCGCAACCTGACCAAAATCTTCAACCAAGTTCAGCAATTC 1080  
1081 TCCAAATTTGATGAAATTTGCTCAAGAAAGTGTTCGAAAGTAGGCGATGGTATGCTTT 1140  
1081 TCCAAATTTGATGAAATTTGCTCAAGAAAGTGTTCGAAAGTAGGCGATGGTATGCTTT 1140  
1141 TGAGGAGATGAGTTTCTGCTTATATCCAGCAAGGATCTTTCAGCAAGCAACGAGC 1200  
1141 TGAGGAGATGAGTTTCTGCTTATATCCAGCAAGGATCTTTCAGCAAGCAACGAGC 1200  
1201 AGGCATTTGATGCAAACTGGCCCAAGCAGGTAAGTTTATCTCATAGCTAGGAGCTAAGAA 1260  
1201 AGGCATTTGATGCAAACTGGCCCAAGCAGGTAAGTTTATCTCATAGCTAGGAGCTAAGAA 1260  
1261 AACTGACCTCCCATCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
1261 AACTGACCTCCCATCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
1321 AATTCCACCAAGATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
1321 AATTCCACCAAGATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
1381 CTTGTTGGAACTCAAGATGCTCAAAATGATAAAGTCAAGTATGATGATGATGATGATGATGAT 1440  
1381 CTTGTTGGAACTCAAGATGCTCAAAATGATAAAGTCAAGTATGATGATGATGATGATGATGAT 1440  
1441 TGCCTTTTCTAGCTCCGATGCTCATCCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 1500  
1441 TGCCTTTTCTAGCTCCGATGCTCATCCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 1500  
1501 CTACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560  
1501 CTACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560

1561 TTATATCTTTTGTATCTCTGATATATTAACAGATGATGAGGGGATGCTTATGTAATCCACA 1620  
1561 TTATATCTTTTGTATCTCTGATATATTAACAGATGATGAGGGGATGCTTATGTAATCCACA 1620  
1621 TATGACCCATAGCCATGCTGATTTAAAGATGATGTTTGTCTGAAGCTGAGAGAGCGGACG 1680  
1621 TATGACCCATAGCCATGCTGATTTAAAGATGATGTTTGTCTGAAGCTGAGAGAGCGGACG 1680  
1681 CAGGCTTATGCTTAAAGAGAAAGGTTTGACCCCTCTTTCGACAGACCATCAGGATTCAG 1740  
1681 CAGGCTTATGCTTAAAGAGAAAGGTTTGACCCCTCTTTCGACAGACCATCAGGATTCAG 1740  
1741 AATATCTGAGGCAAAAGGAGCAGAGATCTATCAACCCGCTGAAAGCAGCTAAGAGGT 1800  
1741 AATATCTGAGGCAAAAGGAGCAGAGATCTATCAACCCGCTGAAAGCAGCTAAGAGGT 1800  
1801 GGCATTTGATCTGCTTATGCTTCAATCTTCAATATCTGTAAGTCAAAAACGTTAGTTT 1860  
1801 GGCATTTGATCTGCTTATGCTTCAATCTTCAATATCTGTAAGTCAAAAACGTTAGTTT 1860  
1861 AATCATACCTCATTTATGACCATTAACCATTAACCATTAACCATTAACCATTAACCATTA 1920  
1861 AATCATACCTCATTTATGACCATTAACCATTAACCATTAACCATTAACCATTAACCATTA 1920  
1921 TTATGAGGCACTTAAAGGGGTATATCTCTGAGGATCTTTTGGGACCTGTCAAAGTACTATGT 1980  
1921 TTATGAGGCACTTAAAGGGGTATATCTCTGAGGATCTTTTGGGACCTGTCAAAGTACTATGT 1980  
1981 CGAATATCCAAACGAACTGCTGCTGATTCAGATTAATGTTTTTGGTAACTGAGGACCATGT 2040  
1981 CGAATATCCAAACGAACTGCTGCTGATTCAGATTAATGTTTTTGGTAACTGAGGACCATGT 2040  
2041 TCAAGAAACCAAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100  
2041 TCAAGAAACCAAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100  
2101 ACCTCAGACAGAAACCTTGAGGAGAAACCCCTCGAGAGAGAAACCCGCAAAAGCGAGAA 2160  
2101 ACCTCAGACAGAAACCTTGAGGAGAAACCCCTCGAGAGAGAAACCCGCAAAAGCGAGAA 2160  
2161 ACAGAGTCTCCAAACCAACAGAGAGAAACAGAGAAATCACCAGAGAAATCAGAGAAC 2220  
2161 ACAGAGTCTCCAAACCAACAGAGAGAAACAGAGAAATCACCAGAGAAATCAGAGAAC 2220  
2221 TCAGGTCGAGACTGAAAGGTTGAGAGAAACCTGAGAGGCTGAGAGATTTACTTGGAAA 2280  
2221 TCAGGTCGAGACTGAAAGGTTGAGAGAAACCTGAGAGGCTGAGAGATTTACTTGGAAA 2280  
2281 AATCCAGGAT 2290  
2281 AATCCAGGAT 2290

RESULT 2  
US-09-769-787-206  
; Sequence 206, Application US/09769787  
; Publication No. US20030091577A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbial Technics Limited  
; APPLICANT: Gilbert, Christophe FG  
; APPLICANT: Hansbro, Philip M  
; TITLE OF INVENTION: Proteins  
; FILE REFERENCE: PWC/P21129W  
; CURRENT APPLICATION NUMBER: US/09/769,787  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: GB 9816337.1  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: US 60/125164  
; PRIOR FILING DATE: 1999-03-19  
; NUMBER OF SEQ ID NOS: 388  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 206  
; LENGTH: 2481

; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-09-769-787-206

Query Match 98.0%; Score 22.4.8; DB 10; Length 2481;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 2268; Conservative 0; Mismatches 22; Indels 3; Gaps 1;

QY	1	TTGTTCTCTATGAATGGTGGTCCGTCACCAAGCTGGTTCAGGTTTAAAGAAAGAGCTCTAATCGAGT	60	DB	1017	CCITCGTTATGTTGTTCAACCCATTGGGTACAGATTCAAGACCAAGAAACAACCAAGTCCACA	1076
DB	57	TTGTTCTCTATGAGTTGGTGGTCCGTCACCAAGCTGGTTCAGGTTTAAAGAAAGAGCTCTAATCGAGT	116	QY	1018	ATCGACTCCGGAACTAGTCCAAAGTCCGCAACCTGCAACCAATCCTCAACCAAGTCCCAAG	1077
QY	61	TTCTTATATAGATGGTGATCAGGCTGGTCAAAAGGCGAGAAACCTTGACACCAAGATGAAGT	120	DB	1077	ATCGACTCCGGAACTAGTCCAAAGTCCGCAACCTGCAACCAATCCTCAACCAAGTCCCAAG	1136
DB	117	TGCTTATATAGATGGTGATCAGGCTGGTCAAAAGGCGAGAAACCTTGACACCAAGATGAAGT	176	QY	1078	CAATCCAAATTGATGAGAAATTTGGTCAAAAGAGCTGTTTCGAAAAAGTAGGCGATGTTATGT	1137
QY	121	CAGTAAGAGGGGGGATCAACGCCGAAACAAATNGTNATCAAGNTTACGATTTACGGATCAAGTTA	180	DB	1137	CAATCCAAATTGATGAGAAATTTGGTCAAAAGAGCTGTTTCGAAAAAGTAGGCGATGTTATGT	1196
DB	177	CAGTAAGAGGGGGGATCAACGCCGAAACAAATNGTNATCAAGNTTACGATTTACGGATCAAGTTA	236	QY	1138	CTTTGAGGAAATGGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTTCAGCAAGAACAGC	1197
QY	181	TGTGACCTCTCATGGAGACCAATATCATTAATAATGGCAAGGTTCTTTATGATGCCAT	240	DB	1197	CTTTGAGGAAATGGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTTCAGCAAGAACAGC	1256
DB	237	TGTGACCTCTCATGGAGACCAATATCATTAATAATGGCAAGGTTCTTTATGATGCCAT	296	QY	1198	AGCAGCATTTGATAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGAGCTAA	1257
QY	241	CATCAGTGAAGAGCTCCTCATGAAGATCCGAATATACGTTGAAGGATTCAGACATTTGT	300	DB	1257	AGCAGCATTTGATAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGAGCTAA	1316
DB	297	CATCAGTGAAGAGCTCCTCATGAAGATCCGAATATACGTTGAAGGATTCAGACATTTGT	356	QY	1258	GAAAACTGACCTCCCATCTAGTGTATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGC	1317
QY	301	CAATGAATCAAGGGTGGTGTATGTCATTAAGTTAAACGGTAAATACATCTATNTACCTTAA	360	DB	1317	GAAAACTGACCTCCCATCTAGTGTATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGC	1376
DB	357	CAATGAATCAAGGGTGGTGTATGTCATTAAGTTAAACGGTAAATACATCTATNTACCTTAA	416	QY	1318	AGAAATTCACCAAGATTTACTTGTATTAAGGTTCGACAAAGTTTGTGAGGCTTTGGA	1377
QY	361	GGATGACGCTCATCGGATTAATATTCGGCAAAAAGAGATTAACCGTCAGAGCAGGA	420	DB	1377	AGAAATTCACCAAGATTTACTTGTATTAAGGTTCGACAAAGTTTGTGAGGCTTTGGA	1436
DB	417	GGATGACGCTCATCGGATTAATATTCGGCAAAAAGAGATTAACCGTCAGAGCAGGA	476	QY	1378	TAACCTGTTTGGAAAGCTCAAGGATGTCNCAAGTGAATAAGTCAAGTTAGTGANGATAT	1437
QY	421	ACGAGTCATAATCAATCAAGAGC...AGATAATGCTGTGCTCGAGCCAGAGCCCA	477	DB	1437	TAACCTGTTTGGAAAGCTCAAGGATGTCNCAAGTGAATAAGTCAAGTTAGTGANGATAT	1496
DB	477	ACGAGTCATAATCAAGAGC...AGATAATGCTGTGCTCGAGCCAGAGCCCA	536	QY	1438	TCCTTGCTCTTTAGCTCCGATTCCTCATCAGAACCGTTTATAGGAAACCAATGCGCAAT	1497
QY	478	AGGAGGTATACACGATGATGGTATTCCTCAATGCACTGATATCTTGGAGCAC	537	DB	1497	TCCTTGCTCTTTAGCTCCGATTCCTCATCAGAACCGTTTATAGGAAACCAATGCGCAAT	1556
DB	537	AGGAGGTATACACGATGATGGTATTCCTCAATGCACTGATATCTTGGAGCAC	596	QY	1498	TACTACACTGATGATGAGATTCAGTAGCCAAAGTTGGCGAGCAAGTACACAACAGAGA	1557
QY	538	GGGTGATGCTTATATGCTTCCTCAGGCACCAATACCAATTTACATTTCTTAAGATGAGTT	597	DB	1557	TACTACACTGATGATGAGATTCAGTAGCCAAAGTTGGCGAGCAAGTACACAACAGAGA	1616
DB	597	GGGTGATGCTTATATGCTTCCTCAGGCACCAATACCAATTTACATTTCTTAAGATGAGTT	656	QY	1558	CGGTTATATCTTTGATCTCGTGATTAACCAAGTGAATAAGTCAAGTTAGTGANGATAT	1617
QY	598	ATCAGCTAGCGAGTTAGTCTGCTCAGAAACCTATTTGAAATGGGAGCAGGATCTCGTCC	657	DB	1617	CGGTTATATCTTTGATCTCGTGATTAACCAAGTGAATAAGTCAAGTTAGTGANGATAT	1676
DB	657	ATCAGCTAGCGAGTTAGTCTGCTCAGAAACCTATTTGAAATGGGAGCAGGATCTCGTCC	716	QY	1618	ACATATGACCCATAGCCACTGGATTTAAAAAGATAGTTTCTGTAAGCTCAGAGAGCGGC	1677
QY	658	TTCTTCAAGTTCTAGTTATTAATCAAAATCCAGCTCAACCAAGATTTGTCAGAGAACCAAA	717	DB	1677	ACATATGACCCATAGCCACTGGATTTAAAAAGATAGTTTCTGTAAGCTCAGAGAGCGGC	1736
DB	717	TTCTTCAAGTTCTAGTTATTAATCAAAATCCAGCTCAACCAAGATTTGTCAGAGAACCAAA	776	QY	1678	AGCCAGGCTTATGCTTAAAGAGAAAGTTTGAACCTCTTCGACAGACCATCAGGATTC	1737
QY	718	TCTGACTGTCACTCCAACTTATATCAAAATCAAGAGGAAACCAATTTCAAGCCCTTTACG	777	DB	1737	AGCCAGGCTTATGCTTAAAGAGAAAGTTTGAACCTCTTCGACAGACCATCAGGATTC	1796
DB	777	TCTGACTGTCACTCCAACTTATATCAAAATCAAGAGGAAACCAATTTCAAGCCCTTTACG	836	QY	1738	AGGAAATATCTGAGGCAAAAGGAGCAGAAAGCTATCTACAACCGGCTGAAAGCAGCTAAGAA	1797
QY	778	TGAATTTGATGCTAAACCTTTATCAGAAACGCAATGTTGAAATCTGATGGCCCTTATTTGGA	837	DB	1797	AGGAAATATCTGAGGCAAAAGGAGCAGAAAGCTATCTACAACCGGCTGAAAGCAGCTAAGAA	1856
DB	837	TGAATTTGATGCTAAACCTTTATCAGAAACGCAATGTTGAAATCTGATGGCCCTTATTTTGA	896	QY	1798	GGTGCCACTTGTATGCTTCAATCTTCAATATCTGTAAGTCAAAACCGTAG	1857
QY	838	CCCAGCGAAATCAAGTCCGAAACCGCAAGGTTGATGCTCCTCATGTGTAAACCAATTA	897	DB	1857	GGTGCCACTTGTATGCTTCAATCTTCAATATCTGTAAGTCAAAACCGTAG	1916
DB	897	CCCAGCGAAATCAAGTCCGAAACCGCAAGGTTGATGCTCCTCATGTGTAAACCAATTA	956	QY	1858	TTTAAATCATACCTCTTATGACCAATACATCAATCAATTTGAGTGGTTTTCAGCAAGG	1917
QY	898	CCACTTTATCCCTTATGACCAAAATGCTCAATTTGAAAGAAACGAATTTGCTGATTTATTC	957	DB	1917	TTTAAATCATACCTCTTATGACCAATACATCAATCAATTTGAGTGGTTTTCAGCAAGG	1976
DB	957	CCACTTTATCCCTTATGACCAAAATGCTCAATTTGAAAGAAACGAATTTGCTGATTTATTC	1016	QY	1918	CTTTTATGAGGCACTTAAAGGGGTATCTCTTGAAGGATCTTTTGGCGACTGTCAAGTACTA	1977
QY	958	CCTTCGTTATCGTTCAAAACCAATTTGGGTAACAGATTCAGAACCAACCAAGTCCACA	1017	DB	1977	CTTTTATGAGGCACTTAAAGGGGTATCTCTTGAAGGATCTTTTGGCGACTGTCAAGTACTA	2036



Db 2097 TGTTCAGAGAAACAAATAATGGTCAAGCTGATCAAAATCAAAACGAGAAACCAAGCGAGGA 2156  
QY 2098 GAAACCTCAGACAGAGAAACCTGAGGAAGAAACCCCTCGAGAGAGAGAAACCGCAAGCGA 2157  
Db 2157 GAAACCTCAGACAGAGAAACCTGAGGAAGAAACCCCTCGAGAGAGAGAAACCGCAAGCGA 2216  
QY 2158 GAAACAGAGTCTCCAAACCAACAGAGGAAACAGAGGAAACAGAGAAATCAACAGAGAAATCAGAAGA 2217  
Db 2217 GAAACAGAGTCTCCAAACCAACAGAGGAAACAGAGGAAACAGAGAAATCAACAGAGAAATCAGAAGA 2276  
QY 2218 ACCTCAGGTCGAGACTGAAAGAGTTGAAGAAACCTGAGAGAGGCTGAAGATTTACTTGG 2277  
Db 2277 ACCTCAGGTCGAGACTGAAAGAGTTGAAGAAACCTGAGAGAGGCTGAAGATTTACTTGG 2336  
QY 2278 AAAAATCCAGGAT 2290  
Db 2337 AAAAATCCAGGAT 2349

RESULT 3  
US-10-412-862-11  
; Sequence 11, Application US/10412862  
; Publication No. US20040052781A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; FILE REFERENCE: 469201-685  
; CURRENT APPLICATION NUMBER: US/10/412,862  
; CURRENT FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 09/468,656  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patent In Ver. 3.0  
; SEQ ID NO 11  
; LENGTH: 2531  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-10-412-862-11

Query Match 97.0%; Score 2220.2; DB 13; Length 2531;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 2253; Conservative 3; Mismatches 16; Indels 18; Gaps 1;

QY 1 TTGTTCTTAAGAACTTGGTCGTACCAAGCTGGTCAGGTTAAGAAAGAGTCTTAATCGAGT 60  
Db 57 TTGTTCTTAAGAACTTGGTCGTACCAAGCTGGTCAGGTTAAGAAAGAGTCTTAATCGAGT 116  
QY 61 TTCTTATATAGATGGTATGATCAGCTGGTCAAAAGGAGAGAACTTACACCAAGATCGAGT 120  
Db 117 TCGTTATATAGATGGTATGATCAGCTGGTCAAAAGGAGAGAACTTACACCAAGATCGAGT 176  
QY 121 CAGTAAGAGGGAGGAGTCAACCGCGAACAATNGTATCAAGATTTACGGATCAAGGTTA 180  
Db 177 CAGTAAGAGGGAGGAGTCAACCGCGAACAATNGTATCAAGATTTACGGATCAAGGTTA 236  
QY 181 TGTGACCTCTCATGGAGACCATATATCAATTAATATGGCAGAGTTCCCTTATGATGCCAT 240  
Db 237 TGTGACCTCTCATGGAGACCATATATCAATTAATATGGCAGAGTTCCCTTATGATGCCAT 296  
QY 241 CATCAGTGAAGAGTCTCTCATGAAAGATCAGATATCATAGTTGAGGATTCAGACATGCT 300  
Db 297 CATCAGTGAAGAGTCTCTCATGAAAGATCAGATATCATAGTTGAGGATTCAGACATGCT 356  
QY 301 CAATGAAATCAAGGGTGGTATGATCAATTAAGGTAAACGGTAAATATCATGTTTACCTTAA 360  
Db 357 CAATGAAATCAAGGGTGGTATGATCAATTAAGGTAAACGGTAAATATCATGTTTACCTTAA 416  
QY 361 GGATGACGCTCATGGGATTAATTCGGAATAAAGAGAGATTAACCGTCAGAGAGCGA 420

Db 417 GGATGCRGCTCATGGGATAATATTCGGACAAAGAGAGATTAACCGTCAGAGCAGGA 476  
QY 421 AGCAGTCAATATCATTAATCACTCAAGAGCAGATATGCTGTGTGTCAGAGCAGAGCCCAAGG 480  
Db 477 AGCAGTCAATATCATTAATCACTCAAGAGCAGATATGCTGTGTGTCAGAGCAGAGCCCAAGG 536  
QY 481 AGCTTATACAAACGGATGATGGGTATATCTTCAATGCATCTGATATCATTAATGAGGACAGGG 540  
Db 537 AGCTTATACAAACGGATGATGGGTATATCTTCAATGCATCTGATATCATTAATGAGGACAGGG 596  
QY 541 TGATGCTTATATCGTTCTCAACGGGAGCAATTAACCATTAATCTTCAAGATGAGTTATC 600  
Db 597 TGATGCTTATATCGTTCTCAACGGGAGCAATTAACCATTAATCTTCAAGATGAGTTATC 656  
QY 601 AGCTAGCGAGTTAGTCTGTCAGAGAGCCTTATGGAATGGAGAGGAGATCTGTCCTTC 660  
Db 657 AGCTAGCGAGTTAGTCTGTCAGAGAGCCTTATGGAATGGAGAGGAGATCTGTCCTTC 716  
QY 661 TTCAAGTCTTAGTTATATGCAAAATCCAGCTCAACCAAGATTTGTAGAGAAACCAATCT 720  
Db 717 TTCAAGTCTTAGTTATATGCAAAATCCAGCTCAACCAAGATTTGTAGAGAAACCAATCT 776  
QY 721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAACATTTCAAGCCTTTTACGTGA 780  
Db 777 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAACATTTCAAGCCTTTTACGTGA 836  
QY 781 ATTGTATGCTAAACCTTATCAGAGGCAATGGAATCTGATGCGCTTATTTGACCC 840  
Db 837 ATTGTATGCTAAACCTTATCAGAGGCAATGGAATCTGATGCGCTTATTTGACCC 896  
QY 841 AGCGCAAAATCAAGTCCGAAACCGCAGAGGTGTAGCTGCTCATGTAACCAATTAACCA 900  
Db 897 AGCGCAAAATCAAGTCCGAAACCGCAGAGGTGTAGCTGCTCATGTAACCAATTAACCA 956  
QY 901 CTTTATCCCTTATGAACAAATGCTGAAATGGAAAAACGAAATGCTGTAATTTATCCCT 960  
Db 957 CTTTATCCCTTATGAACAAATGCTGAAATGGAAAAACGAAATGCTGTAATTTATCCCT 1016  
QY 961 TCGTTATCGTTCAAAACCTTGGTACCAAGATTCAGAGCAGAGCAACCAAGTCCCAATC 1020  
Db 1017 TCGTTATCGTTCAAAACCTTGGTACCAAGATTCAGAGCAGAGCAACCAAGTCCCAATC 1076  
QY 1021 GACTCGGAAACCTAGTCCGAAACCGCAGAGGTGTAGCTGCTCATGTAACCAATTAACCA 1080  
Db 1077 GACTCGGAAACCTAGTCCGAAACCGCAGAGGTGTAGCTGCTCATGTAACCAATTAACCA 1118  
QY 1081 TCCAAATGATGAGAAATTTGGTCAAAAGAGCTGTTGAAAGAGTAGGCGATGGTTATGCTTT 1140  
Db 1119 TCCAAATGATGAGAAATTTGGTCAAAAGAGCTGTTGAAAGAGTAGGCGATGGTTATGCTTT 1178  
QY 1141 TGAGAGATGAGGTTCTCGTTATATCCAGCAGAGATCTTTCAGCAGAGAACAGCAGC 1200  
Db 1179 TGAGAGATGAGGTTCTCGTTATATCCAGCAGAGATCTTTCAGCAGAGAACAGCAGC 1238  
QY 1201 AGGCATTGATAGCAAACTGGCCAGCAGAGAAATTTATCTCAATAAGCTAGGAGCTAAGAA 1260  
Db 1239 AGGCATTGATAGCAAACTGGCCAGCAGAGAAATTTATCTCATAGCTAGGAGCTAAGAA 1298  
QY 1261 AACTGACCTCCATCTAGTATCGAGAAATTTTAAATAGGCTTATGATCTTACTAGCAAG 1320  
Db 1299 AACTGACCTCCATCTAGTATCGAGAAATTTTAAATAGGCTTATGATCTTACTAGCAAG 1358  
QY 1321 AATTCAACCAAGATTTACTTGAATAAAGTTCGAACTGATGTTGAGGCTTTCGATTA 1380  
Db 1359 AATTCAACCAAGATTTACTTGAATAAAGTTCGAACTGATGTTGAGGCTTTCGATTA 1418  
QY 1381 CCGTTGGAACGACTCAAGGATGTCNCAAGTATATAAGTCAAGTTAGTGGAGATATTCT 1440  
Db 1419 CCGTTGGAACGACTCAAGGATGTCNCAAGTATATAAGTCAAGTTAGTGGAGATATTCT 1478  
QY 1441 TGCCTTCTTAGTCCGATTCGTTATCCAGAGAGTTTGAAGAAACCAATGCGCAATTTAC 1500

1479	Db	TGCTCTTTAGCTCCGATTTCGTATCAACGAAGTTTTAGGAAAAACAATCGCAATTAC	1538
1501	QY	CTACACTGATGATGAGANTCAAGTAGCCAGTGGCAGGCAGCTACACAAGAAGACGG	1560
1539	Db	CTACACTGATGATGAGATTCAAGTAGCCAGTGGCAGGCAGCTACACAAGAAGACGG	1598
1561	QY	TTATATCTTTTGANCTCGTGATATAACCCTGATGAGGGGGATGCCCTATGTAACTCCACA	1620
1599	Db	TTATATCTTTTGATCCTCGTGATATAACCCTGATGAGGGGGATGCCCTATGTAACTCCACA	1658
1621	QY	TATGACCCATAGCCACTCGGATTAAAATAATAGTTTTGCTGAAGCTCGAGAGCGCGCAGC	1680
1659	Db	TATGACCCATAGCCACTCGGATTAAAATAATAGTTTTGCTGAAGCTCGAGAGCGCGCAGC	1718
1681	QY	CCAGGCTTATGCTTAAGAGAAAAAGTTTGAAGCCCTCCTTCGACAGACCATCAGATTTCAGG	1740
1719	Db	CCAGGCTTATGCTTAAGAGAAAAAGTTTGAAGCCCTCCTTCGACAGACCATCAGATTTCAGG	1778
1741	QY	AATATCTGAGGCAAAAGAGCAGAAAGCTTCTACAACCGCGTCAAGACGACTAAGAGGT	1800
1779	Db	AATATCTGAGGCAAAAGAGCAGAAAGCTTCTACAACCGMGTTAAGACGACTAAGAGGT	1838
1801	QY	GCCACTTGATCGTATGCTTTACAATCTTCAATATCTGTAGAAAGTCAAAAACGGTAGTTT	1860
1839	Db	GCCACTTGATCGTATGCTTTACAATCTTCAATATCTGTAGAAAGTCAAAAACGGTAGTTT	1898
1861	QY	AATCATACTCATTTATGACCATTTACATACATCAATTTGAGTGGTTTGACGAAGGCCT	1920
1899	Db	AATCATACCTTCATTTATGACCATTTACCATACATCAAATTTGAGTGGTTTGACGAAGGCCT	1958
1921	QY	TTATGAGGCACCTAAAGGGGTATCTCTTCAGGATCTTTTCGCGACATGTCGAAGTACTATGT	1980
1959	Db	TTATGAGGCACCTPAAGGGGTATCTCTTCAGGATCTTTTCGCGACATGTCGAAGTACTATGT	2018
1981	QY	CGAACATCCAACGGAACGTCGCAATTCATAATATGGTTTTGTTAAGCTAGCACCATGT	2040
2019	Db	CGAACATCCAACGGAACGTCGCAATTCATAATATGGTTTTGTTAAGCTAGCACCATGT	2078
2041	QY	TCBAAGAGACAAAATGTTCAAGCTGATCCATCAAAACGAAAAACCAAGCGAGAGAA	2100
2079	Db	TCBAAGAGACAAAATGTTCAAGCTGATCCATCAAAACGAAAAACCAAGCGAGAGAA	2138
2101	QY	ACCTCAGACAGAAAAACCTTGAGGAAGAAACCCCTCGAGAAAGAAAAACCGCAAGCGAGAA	2160
2139	Db	ACCTCAGACAGAAAAACCTTGAGGAAGAAACCCCTCGAGAAAGAAAAACCGCAAGCGAGAA	2198
2161	QY	ACCAGAGTCTCAAAAACCAACAGAGGAAACAGAAAGATCAACAGAGGAATCAGAAAGCC	2220
2199	Db	ACCAGAGTCTCAAAAACCAACAGAGGAAACAGAAAGATCACCAGAGGAATCAGAAAGCC	2258
2221	QY	TCAAGTTCGAGACTGAAAAAGTTCAAGAAACACTGAGAGAGGCTGAAGATTTACTTCGAAA	2280
2259	Db	TCAGGTCGAGACTGAAAAAGTTCAAGAAACACTGAGAGAGGCTGAAGATTTACTTCGAAA	2318
2281	QY	AATCCAGGAT	2290
2319	Db	AATCCAGGAT	2328

RESULT 4  
US-10-412-850-11  
; Sequence 11, Application US/10412850  
; Publication No. US20040001836A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; TITLE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-686  
; CURRENT APPLICATION NUMBER: US/10/412,850  
; CURRENT FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 09/468,656

```

; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 11
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-412-850-11

```

Query Match	97.0%	Score 2220.2	DB 16	Length 2531
Best Local Similarity	98.4%	Pred. No. 0		
Matches 2253	Conservative 3	Mismatches 16	Indels 18	Gaps 1
QY	1	TTCTTCTATGAACCTGGTCGTCCACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCCAGT	60	
Db	57	TTCTTCTATGAGCTTTGGACGTTTACCAGCTGGTCAGGATAAGAAAGAGTCTAATCCAGT	116	
QY	61	TTCTTTATATAGATGGTGATCAGGCTGGTCAAAAAGGCAGAAAACTTGACACCAAGTGAAGT	120	
Db	117	TGCTTATATAGATGGTGATCAGGCTGGTCAAAAGGCAGAAAATTGACACCAAGTGAAGT	176	
QY	121	CAGTAAGAGGAGGGGATCAACGCCGAAACAAATNGTNATCAAGATTACGATCAAGGTTA	180	
Db	177	CAGTAAGAGGAGGGGATCAACGCCGAAACAAATTTGTTATCAAGATTACGATCAAGGTTA	236	
QY	181	TGTGACCTCTCATGGAGACCATTTATCATTTACTATATAATGCGCAAGTTCCTTTATGATGCCAT	240	
Db	237	TGTGACCTCTCATGGAGACCATTTATCATTTACTATATAATGCGCAAGTTCCTTTATGATGCCAT	296	
QY	241	CATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTTATCAGTTGAAGGATTCAGACATTGT	300	
Db	297	CATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTTATCAGTTGAAGGATTCAGACATTGT	356	
QY	301	CAATGAAATCAAGGGTGGTTATGTCATTAAGGTAAACGGTAAATACTATGNTACCTTAA	360	
Db	357	CAATGAAATCAAGGGTGGTTATGTCATTAAGGTAAACGGTAAATACTATGTTACCTTAA	416	
QY	361	GGATGCAGCTCATGCCGATAAATTCGGAACAAAGAGAGATTAAACGTCAGAAAGCAGGA	420	
Db	417	GGATGCRGCTCATGCCGATAAATTCGGAACAAAGAGAGATTAAACGTCAGAAAGCAGGA	476	
QY	421	AGCGATCTATATCATPAATCTCAAGAGCAGATAATGCTGTGTCGACCCAGAGCCCAAGG	480	
Db	477	AGCGATCTAATATCAATACTCAAGAGCAGATAATGCTGTGTCGACCCAGAGCCCAAGG	536	
QY	481	ACGTTATACAAAGGATGAGGGGTATATCTTCAATGCATCTGATATCATTCAGGACACGGG	540	
Db	537	ACGTTATACAAAGGATGAGGGGTATATCTTCAATGCATCTGATATCATTCAGGACACGGG	596	
QY	541	TGATGCTTATATGTTTCTTCACGGCGACCATTAACCATTAATTCCTAAGAAATCAGTTATC	600	
Db	597	TGATGCTTATATGTTTCTTCACGGCGACCATTAACCATTAATTCCTAAGAAATCAGTTATC	656	
QY	601	AGCTAGCGAGTTAGTCTGTCAGAACCTTATGGAAATGGGAAGCAGGGATCTCGTCCTTC	660	
Db	657	AGCTAGCGAGTTAGTCTGTCAGAACCTTATGGAAATGGGAAGCAGGGATCTCGTCCTTC	716	
QY	661	TTCAAGTTCTAGTTTATATGCAAAATCCAGCTCAACCAAGATTGTCAGAGAACCACAATCT	720	
Db	717	TTCAAGTTCTAGTTTATATGCAAAATCCAGCTCAACCAAGATTGTCAGAGAACCACAATCT	776	
QY	721	GACTGTCACTCCAACCTTATCATCAAAATCAAGGGGAAAAACATTTCAAGCCTTTTACGTGA	780	
Db	777	GACTGTCACTCCAACCTTATCATCAAAATCAAGGGGAAAAACATTTCAAGCCTTTTACGTGA	836	
QY	781	ATTGTATGCTAAACCCCTTATCAGAACCGCATGTGGAAATCTGATGGCCTTATTTTCACCC	840	
Db	837	ATTGTATGCTAAACCCCTTATCAGAACCGCATGTGGAAATCTGATGGCCTTATTTTCACCC	896	
QY	841	AGCGCAAAATCAAGTTCGAAACCGCAGAGGTGTAGCTGTCCCTCATGTGTAACCATTTACCA	900	

897 AGCGCAATCAAGTCGAAACCGCCAGAGGTTAGCTGTCCCTCATGGTAAACCATACCA 956  
901 CTTTATCCCTTATGAACAAATGTCTGAATGGAAAAACGAATTCGTCTGATTTATCCCT 960  
957 CTTTATCCCTTATGAACAAATGTCTGAATGGAAAAACGAATTCGTCTGATTTATCCCT 1016  
961 TCGTTATCGTTCAAAACCAATGGGTACAGATTCAAGACGAGAAACCAAGTCCACAATC 1020  
1017 TCGTTATCGTTCAAAACCAATGGGTACAGATTCAAGACGAGAAACCAAGTCCACAATC 1076  
1021 GACTCGGAACTAGTCCCAAGTCCGCAACCTGCACCAATCTCAACCACTCAAGCAA 1080  
1077 GACTCCGAACCTAGTCCCAAGTCCG-----CAACCACTCAAGCAA 1118  
1081 TCCAAATGATGAGAAATTCGTCAAAGAAAGCTGTCGAAAGTAGGCGATGTTATGTCCT 1140  
1119 TCCAAATGATGAGAAATTCGTCAAAGAAAGCTGTCGAAAGTAGGCGATGTTATGTCCT 1178  
1141 TGAGGAAATGAGGTTTCGTTATATCCAGCAAGGATCTTTTCAGCAGAAACAGCAGC 1200  
1179 TGAGGAAATGAGGTTTCGTTATATCCAGCAAGGATCTTTTCAGCAGAAACAGCAGC 1238  
1201 AGGCATTGATAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAAGCTAGGAGCTAAGAA 1260  
1239 AGGCATTGATAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAAGCTAGGAACTAAGAA 1298  
1261 AACTGACCTCCCATCTAGTGATCGAGAAATTTACATAAAGGCTTATGACTTACTAGCAAG 1320  
1299 AACTGACCTCCCATCTAGTGATCGAGAAATTTACATAAAGGCTTATGACTTACTAGCAAG 1358  
1321 AACTCACAAGATTTACTTGATTAATAAAGGTCGACAAAGTTGATTTTGGAGCTTTGGATAA 1380  
1359 AACTCACAAGATTTACTTGATTAATAAAGGTCGACAAAGTTGATTTTGGAGCTTTGGATAA 1418  
1381 CTTGTGGAACGACTCAAGGATGTCNCAAGTGAAGAATCAAGTCAAGTCAAGTCAAGTCAAGT 1440  
1419 CTTGTGGAACGACTCAAGGATGTCNCAAGTGAAGAATCAAGTCAAGTCAAGTCAAGTCAAGT 1478  
1441 TGCCTCTTCTAGTCCGATTCGTCATCAGAAAGTTTGGAAACCAAAATGCGCAAAATTC 1500  
1479 TGCCTCTTCTAGTCCGATTCGTCATCAGAAAGTTTGGAAACCAAAATGCGCAAAATTC 1538  
1501 CTACACTGATGATGATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1560  
1539 CTACACTGATGATGATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1598  
1561 TTATATCTTTGATCTCGTGATATACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1620  
1599 TTATATCTTTGATCTCGTGATATACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1658  
1621 TATGACCATAGCCACTGATTTAAAGAAAGATGTTGCTGAGCTGAGAGCGGCGC 1680  
1659 TATGACCATAGCCACTGATTTAAAGAAAGATGTTGCTGAGCTGAGAGCGGCGC 1718  
1681 CCAGGCTTATGCTAAAGAAAGGTTTGACCCCTCTCTCGACAGACCATCAGATTCAGG 1740  
1719 CCAGGCTTATGCTAAAGAAAGGTTTGACCCCTCTCTCGACAGACCATCAGATTCAGG 1778  
1741 AAATCTGAGGCAAAAGGAGCAGAGCTATCTACAAACCGCGTGAAGCAGCTAAGAGGT 1800  
1779 AAATCTGAGGCAAAAGGAGCAGAGCTATCTACAAACCGCGTGAAGCAGCTAAGAGGT 1838  
1801 GCCACTTCATCGTATGCTTCAAACTTCAATATCTAGAGTCAAGAAACCGGTAGTTT 1860  
1839 GCCACTTCATCGTATGCTTCAAACTTCAATATCTAGAGTCAAGAAACCGGTAGTTT 1898  
1861 AATCATACCTCATTTAGCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1920  
1899 AATCATACCTCATTTAGCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1958  
1921 TTATGAGGCACTAAGGGGTATCTCTTGGAGATCTTTTGGGCGAGTGTCAAGTACTATGT 1980  
1959 TTATGAGGCACTAAGGGGTATCTCTTGGAGATCTTTTGGGCGAGTGTCAAGTACTATGT 2018

QY 1981 CGAAATCCAAACGAACTCGCATTCAGATAATGTTTTGTTAAAGCTAGCGACCATGT 2040  
Db 2019 CGAAATCCAAACGAACTCGCATTCAGATAATGTTTTGTTAAAGCTAGCGACCATGT 2078  
QY 2041 TCAAGAAACAAATAATGCTCAAGCTGATACCAATCAAAACGAAACCAAGCGAGAGAA 2100  
Db 2079 TCAAGAAACAAATAATGCTCAAGCTGATACCAATCAAAACGAAACCAAGCGAGAGAA 2138  
QY 2101 ACCTCAGACAGAAACCTGAGGAGAAACCTCTCGAGAGAAACCGCAAGCGAGAA 2160  
Db 2139 ACCTCAGACAGAAACCTGAGGAGAAACCTCTCGAGAGAAACCGCAAGCGAGAA 2198  
QY 2161 ACCAGAGTCTCAAAACCAACAGAGAGAACCAAGAAATCACCAGAGGAATCAGAGAAC 2220  
Db 2199 ACCAGAGTCTCAAAACCAACAGAGAGAACCAAGAGAAATCACCAGAGGAATCAGAGAAC 2258  
QY 2221 TCAGGTCGAGACTGAAAGGTTGAAAGAAACCTGAGAGAGGCTGAAGATTTACTTTGAA 2280  
Db 2259 TCAGGTCGAGACTGAAAGGTTGAAAGAAACCTGAGAGAGGCTGAAGATTTACTTTGAA 2318  
QY 2281 AATCCAGGAT 2290  
Db 2319 AATCCAGGAT 2328

RESULT 5  
US-10-387-783-11  
; Sequence 11, Application US/10387783  
; Publication No. US2004005331A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; FILE REFERENCE: 469201-683  
; CURRENT APPLICATION NUMBER: US/10/387,783  
; PRIOR FILING DATE: 2003-03-13  
; PRIOR APPLICATION NUMBER: 09/468,656  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patent in Ver. 3.0  
; SEQ ID NO 11  
; LENGTH: 2531  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-10-387-783-11

Query Match 97.0%; Score 2220.2; DB 16; Length 2531;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 2253; Conservative 3; Mismatches 16; Indels 18; Gaps 1;  
QY 1 TTGTTCTTATGAACTTGGTCTGCTCACCAGCTGTCAGGTTAAGAAAGAGTCTAATCGAGT 60  
Db 57 TTGTTCTTATGAGCTTGGAGCTTTACCAAGCTGTCAGGATAAGAAAGAGTCTAATCGAGT 116  
QY 61 TTCTTATATAGATGGTGATCAGGCTGCTCAAAAGCGCAGAAACCTTGACACAGATGAAGT 120  
Db 117 TGCTTATATAGATGGTGATCAGGCTGCTCAAAAGCGCAGAAACCTTGACACAGATGAAGT 176  
QY 121 CAGTAAGAGAGGAGGATCAACCGCAAAATNGTATCAAGATTAAGGATCAAGGTTA 180  
Db 177 CAGTAAGAGAGGAGGATCAACCGCAAAATNGTATCAAGATTAAGGATCAAGGTTA 236  
QY 181 TGTGAGCTCTATGAGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240  
Db 237 TGTGAGCTCTATGAGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 296  
QY 241 CATCAGTGAAGAGCTCTCTCATGAAGATCCGAATTAATTAATTAATTAATTAATTAATTAAT 300

Db 297 CACAGTGAAGAGCTCCTCATGAAGATCCGAATTATCAGTTGAGGATTCAGCAATTGT 356  
QY 301 CAATGAATCAAGGGTGGTTATGTCAATTAAGGTAAACGGTAAATACTATGNTACTTTAA 360  
Db 357 CAATGAATCAAGGGTGGTTATGTCAATTAAGGTAAACGGTAAATACTATGNTACTTTAA 416  
QY 361 GGATGAGCTCATCGGATATATTCGACAAAGAGAGATTAAACGTCAAGACGAGA 420  
Db 417 GGATGCGCTCATCGGATATATTCGACAAAGAGAGATTAAACGTCAAGACGAGA 476  
QY 421 ACGCAGTCATAATCATTAACCAAGAGCAGATAATGCTGTGTCAGCCAGAGCCCAAGG 480  
Db 477 ACGCAGTCATAATCATTAACCAAGAGCAGATAATGCTGTGTCAGCCAGAGCCCAAGG 536  
QY 481 ACGTTATACAAACGGATGATGGGTATATCTCAATGCAATCTGATATCAATTAAGATGATTC 600  
Db 537 ACGTTATACAAACGGATGATGGGTATATCTCAATGCAATCTGATATCAATTAAGATGATTC 656  
QY 541 TGATGCTTATATCGTTCCTCAACGGGACCATTAACATTAATTCCTAAGATGATTC 600  
Db 597 TGATGCTTATATCGTTCCTCAACGGGACCATTAACATTAATTCCTAAGATGATTC 656  
QY 601 AGCTAGCGAGTTAGCTGTCGAGAGCCCTATTTGGAATGGGAGCAGGGATCTCGTCCCTTC 660  
Db 657 AGCTAGCGAGTTAGCTGTCGAGAGCCCTATTTGGAATGGGAGCAGGGATCTCGTCCCTTC 716  
QY 661 TTCAAGTTCTAGTTATATGTAATCAAGTCAACCAAGATTTGTCAGAGAACCAAAATCT 720  
Db 717 TTCAAGTTCTAGTTATATGTAATCAAGTCAACCAAGATTTGTCAGAGAACCAAAATCT 776  
QY 721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACAAATTTCAAGCCTTTTACGTGA 780  
Db 777 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACAAATTTCAAGCCTTTTACGTGA 836  
QY 781 ATTGTATGCTTAAACCTTTATCAGAACCCATGTCGAATCTGATGCGCTTATTTTCGACC 840  
Db 837 ATTGTATGCTTAAACCTTTATCAGAACCCATGTCGAATCTGATGCGCTTATTTTCGACC 896  
QY 841 AGCGCAATCACAAGTCAAGCCGACAGAGGTGATGTCCTCATGTGTAACCATTAACA 900  
Db 897 AGCGCAATCACAAGTCAAGCCGACAGAGGTGATGTCCTCATGTGTAACCATTAACA 956  
QY 901 CTTTATCCCTTATGAAACAAATGCTGAATGGAATAACGAATTCGTATTTATTTCCCT 960  
Db 957 CTTTATCCCTTATGAAACAAATGCTGAATGGAATAACGAATTCGTATTTATTTCCCT 1016  
QY 961 TCGTTATCGTTCAAAACCATTTGGGTACAGATTTCAAGACCAGAACCAACCAAGTCCCAATC 1020  
Db 1017 TCGTTATCGTTCAAAACCATTTGGGTACAGATTTCAAGACCAGAACCAACCAAGTCCCAATC 1076  
QY 1021 GACTCCGGAACCTAGTCCAGTCCGCAACCTGCACCAATCTCAACCCAGCTCCCAAGCAA 1080  
Db 1077 GACTCCGGAACCTAGTCCAGTCCG-----CAACCCAGCTCCAGCAA 1118  
QY 1081 TCAATTTGATGAGAAATGTTCAAGAGCTGTTTCGAAAGTGAAGCGATGTTATGCTT 1140  
Db 1119 TCAATTTGATGAGAAATGTTTCGAAAGCTGTTTCGAAAGTGAAGCGATGTTATGCTT 1178  
QY 1141 TGAGGAGATGAGATTTCTCGTTATATCCCGACAGAGATCTTTCCAGCAGAAACAGCAGC 1200  
Db 1179 TGAGGAGATGAGATTTCTCGTTATATCCCGACAGAGATCTTTCCAGCAGAAACAGCAGC 1238  
QY 1201 AGGCATTGATAGCAAACTGGCCCAAGCAGGAAGTTTATCTCATTAAGCTAGGAGCTTAAGAA 1260  
Db 1239 AGGCATTGATAGCAAACTGGCCCAAGCAGGAAGTTTATCTCATTAAGCTAGGAACTAAGAA 1298  
QY 1261 AACTGACCTCCCACTAGTGAATGAGAAATTTTACAATAAGGCTTTATGACTTTACTAGCAAG 1320  
Db 1299 AACTGACCTCCCACTAGTGAATGAGAAATTTTACAATAAGGCTTTATGACTTTACTAGCAAG 1358  
QY 1321 AATTCAACAGATTTACTTGATATAAAGGTGCGCAAGTTGATTTTGAAGGCTTTGGATAA 1380  
Db 1359 AATTCAACAGATTTACTTGATATAAAGGTGCGCAAGTTGATTTTGAAGGCTTTGGATAA 1418

QY 1381 CCTGTTGGAACGACTCAAGGATGTCNCAAGTGTATAAAGTCAAGTTAGTGGANGATATTTCT 1440  
Db 1419 CCTGTTGGAACGACTCAAGGATGTCNCAAGTGTATAAAGTCAAGTTAGTGGANGATATTTCT 1478  
QY 1441 TGCCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTAGHAAACCAAAATCGCNAATATAC 1500  
Db 1479 TGCCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTAGHAAACCAAAATCGCNAATATAC 1538  
QY 1501 CTACACTGATGATGAGATTCAAGTAGCTGAGGAGCAAGTACACACAGAAAGACGG 1560  
Db 1539 CTACACTGATGATGAGATTCAAGTAGCTGAGGAGCAAGTACACAGAAAGACGG 1598  
QY 1561 TTATATCTTTGATCTCGTGATATAACCAAGTGTAGGGGGATGCCCTATGTAACTCCACA 1620  
Db 1599 TTATATCTTTGATCTCGTGATATAACCAAGTGTAGGGGGATGCCCTATGTAACTCCACA 1658  
QY 1621 TATGACCCATAGCCATCGGATTAATAAAGATAGTTTGTCTGAAGCTGAGAGACGGCAGC 1680  
Db 1659 TATGACCCATAGCCATCGGATTAATAAAGATAGTTTGTCTGAAGCTGAGAGACGGCAGC 1718  
QY 1681 CCAGGCTTATGCTTAAAGAGAAAGGTTTGACCCCTCTTCGACAGACCATCAGGATTCAGG 1740  
Db 1719 CCAGGCTTATGCTTAAAGAGAAAGGTTTGACCCCTCTTCGACAGACCATCAGGATTCAGG 1778  
QY 1741 AAATACCTGAGGCAAAAGAGAGCAAGCTATCTAACAACCGCTGAAAGCAGCTAAGAAAGGT 1800  
Db 1779 AAATACCTGAGGCAAAAGAGAGCAAGCTATCTAACAACCGMTGAAAGCAGCTAAGAAAGGT 1838  
QY 1801 GGCATTGATGATGCTTACCTTAATCTTCAATATATCTGTAGAAGTCAAAAACGGTAGTTT 1860  
Db 1839 GGCATTGATGATGCTTACCTTAATCTTCAATATATCTGTAGAAGTCAAAAACGGTAGTTT 1898  
QY 1861 AATCATACCTTATGATGACCATTAACCATTAACCAATTTAGTGGTTCACGAGGCCT 1920  
Db 1899 AATCATACCTTATGATGACCATTAACCATTAACCAATTTAGTGGTTCACGAGGCCT 1958  
QY 1921 TTATGAGGCACCTAAGGGGTATATCTTTCAGGATCTTTTGGCGACTGTCAAGTACTATGT 1980  
Db 1959 TTATGAGGCACCTAAGGGGTATATCTTTCAGGATCTTTTGGCGACTGTCAAGTACTATGT 2018  
QY 1981 CGAAATCAAAACGAAACGTCGCAATTCAGATTAATGTTTGGTAAACGCTAGCGACCATGT 2040  
Db 2019 CGAAATCAAAACGAAACGTCGCAATTCAGATTAATGTTTGGTAAACGCTAGCGACCATGT 2078  
QY 2041 TCAAGAAACAAATGTTCAAGCTGATACCATTAACCGAAGAACCAAGACGAGAGAA 2100  
Db 2079 TCAAGAAACAAATGTTCAAGCTGATACCATTAACCGAAGAACCAAGACGAGAGAA 2138  
QY 2101 ACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAGAGAAACCGCAAGCGAGAA 2160  
Db 2139 ACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAGAGAAACCGCAAGCGAGAA 2198  
QY 2161 ACCAGAGTCTCAAAACCAACAGAGAAACCAAGAAATCACCAGAGAAATCAGAAAGAAC 2220  
Db 2199 ACCAGAGTCTCAAAACCAACAGAGAAACCAAGAAATCACCAGAGAAATCAGAAAGAAC 2258  
QY 2221 TCAGTTCGAGACTGAAAGGTTGAAGAAAACTGAGAGAGGCTGAAGATTTACTTGGAAA 2280  
Db 2259 TCAGTTCGAGACTGAAAGGTTGAAGAAAACTGAGAGAGGCTGAAGATTTACTTGGAAA 2318  
QY 2281 AATCCAGGAT 2290  
Db 2319 AATCCAGGAT 2328

## RESULT 6

US-09-884-465A-5  
; Sequence 5, Application US/09884465A  
; Publication No. US20030077293A1  
; GENERAL INFORMATION:  
; APPLICANT: Shire Biochem, Inc.  
; APPLICANT: Hamel, Josee

APPLICANT: Brodeur, Bernard  
APPLICANT: Martin, Denis  
APPLICANT: Charland, Nathalie  
APPLICANT: Ouellet, Catherine  
TITLE OF INVENTION: Streptococcus Antigens  
FILE REFERENCE: 055190-0044  
CURRENT APPLICATION NUMBER: US/09/884,465A  
CURRENT FILING DATE: 2001-06-20  
PRIOR APPLICATION NUMBER: 60/212,683  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 384  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5

LENGTH: 2639

TYPE: DNA

ORGANISM: Streptococcus pneumoniae

US-09-884-465A-5

Query Match 88.3%; Score 2022.2; DB 10; Length 2639;  
Best Local Similarity 92.6%; Pred. No. 0;  
Matches 2142; Conservative 0; Mismatches 148; Indels 24; Gaps 1;

1 TTGTTCCCTATGAACCTTGGTCGTCACCAAGCTGTCAGGTTTAAAGAGTCTTAATCGAGT 60  
170 TTGTTCCCTATGAACCTTGGTCGTCACCAAGCTGTCAGGTTTAAAGAGTCTTAATCGAGT 229

61 TTCTTATATAGATGGTGTATCAGGCTGGTCAAAAGGCAGAAAACTTGACACCCAGATGAAGT 120  
230 TTCTTATATAGATGGTGTATCAGGCTGGTCAAAAGGCAGAAAACTTGACACCCAGATGAAGT 289

121 CAGTAAGAGGAGGGGGATCAACCGCGAAACAAATNGTNATCAAGATTAACGGATCAAGGTTA 180  
290 CAGTAAGAGAGAGGGGGATCAACCGCGAAACAAATNGTNATCAAGATTAACGGATCAAGGTTA 349

181 TGTGACCTCTCATGAGACCCATTAATCAATTAATTAATGCGAAGTTCCTTATGATGCCAT 240  
350 TGTGACCTCTCATGAGACCCATTAATCAATTAATTAATGCGAAGTTCCTTATGATGCCAT 409

241 CATCAGTGAAGAGCTTCCTCATGAAAGATCCGAATTAATCAGTTGAAGGATTCAGACATTTG 300  
410 CATCAGTGAAGAGCTTCCTCATGAAAGATCCGAATTAATCAGTTGAAGGATTCAGACATTTG 469

301 CAATGAATCAAGGGTGGTTATGTCATTAAGGTAAACCGTAAATATCATGTGTTACCTTAA 360  
470 CAATGAATCAAGGGTGGTTATGTCATTAAGGTAAACCGTAAATATCATGTGTTACCTTAA 529

361 GGATGCGAGCTCATGCGGATATATTCGGCAAAAGAGAGATTAACGTCAGAGCAGGA 420  
530 AGATGCGGCGCCATGCGGCAATATTCGGCAAAAGAGAGATTAACGTCAGAGCAGGA 589

421 ACGCAGTCATATCAATACTCAAGAGCAGATTAATGCTGTTGCTGAGCCAGAGCCCAAGG 480  
590 ACA CAGTCATATCAATACTCAAGAGCAGATTAATGCTGTTGCTGAGCCAGAGCCCAAGG 649

481 ACGTTATACAACGAGATGGGTATATCTTCAATGCATCTGATATCAATTAAGGACACGGG 540  
650 ACGTTATACAACGAGATGGGTATATCTTCAATGCATCTGATATCAATTAAGGACACGGG 709

541 TGATGCTTATATCTTCTCAGCGGACCATTAACATTAACATTCCTTAAGATGAGTTATC 600  
710 TGATGCTTATATCTTCTCAGCGGACCATTAACATTAACATTCCTTAAGATGAGTTATC 769

601 AGTAGCCAGTTAGTGTCTGTCAGAGGCTTATGGAATGGAGAGCAGGATCTCGTCTTTC 660  
770 AGTAGCCAGTTAGTGTCTGTCAGAGGCTTATGGAATGGAGAGCAGGATCTCGTCTTTC 829

661 TTCAAGTCTTAGTTAATATGCAAAATCCAGCTCAACCAAGATTTGTAGAGAACCAAACTCT 720  
830 TTCAAGTCTTAGTTAATATGCAAAATCCAGTTCACCAAGATTTGTAGAGAACCAAACTCT 889

721 GACTGTCTACTCAACTTATCAATAATCAAGGGGAAACATTTCAAGCTTTTACGTGA 780  
890 GACTGTCTACTCAACTTATCAATAATCAAGGGGAAACATTTCAAGCTTTTACGTGA 949

QY 781 ATTGATGCTAAACCCCTTATCAGAACGCCATGTGAAATCTGATGCGCTTATTTTCGACC 840  
DB 950 ATTGATGCTAAACCCCTTATCAGAACGCCATGTGAAATCTGATGCGCTTATTTTCGACC 1009

QY 841 AGCGCAATCACAAGTGCAGACCGCCAGAGGTGTAGCTGCTCTCATGTGTAACCATACCA 900  
DB 1010 AGCGCAATCACAAGTGCAGACCGCCAGAGGTGTAGCTGCTCTCATGTGTAACCATACCA 1069

QY 901 CTTTATCCCTTATGAAACAAATGCTGAAATTCGAAAAACGAATTCGCTGATTTATTTCCCT 960  
DB 1070 CTTTATCCCTTATGAAACAAATGCTGAAATTCGAAAAACGAATTCGCTGATTTATTTCCCT 1129

QY 961 TCGTTATTCGTTCAAAACCAATTCGGGTACAGATTCAGACCAAGCAAAACCAAGTCCACAATC 1020  
DB 1130 TCGTTATTCGTTCAAAACCAATTCGGGTACAGATTCAGACCAAGCAAAACCAAGTCCACAATC 1189

QY 1021 GACTCCGGAACTTAGTCCAAAGTCGGGAACTTGACCCAAATCCTCAACAGCTCCAAAGAA 1080  
DB 1190 GACTCCGGAACTTAGTCCAAAGTCGGGAACTTGACCCAAATCCTCAACAGCTCCAAAGAA 1249

QY 1081 TCCAAATGTAGAGAAATTCGTCAAAGAGCTGTTTCGAAAAAGTAGGCGATGTTATGCTTT 1140  
DB 1250 TCCAAATGTAGAGAAATTCGTCAAAGAGCTGTTTCGAAAAAGTAGGCGATGTTATGCTTT 1309

QY 1141 TGAGGGAATTCGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTTCAGCAGAAACAGCAGC 1200  
DB 1310 TGAGGGAATTCGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTTCAGCAGAAACAGCAGC 1369

QY 1201 AGCAATGTAGCAAACTGCGCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTTAAGAA 1260  
DB 1370 AGCAATGTAGCAAACTGCGCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTTAAGAA 1429

QY 1261 AACTGACCTCCCATCTAGTATCGAGAAATTTTACAATAGGCTTATGATTTACTAGCAG 1320  
DB 1430 AACTGACCTCCCATCTAGTATCGAGAAATTTTACAATAGGCTTATGATTTACTAGCAG 1489

QY 1321 AATTCAACCAAGATTTACTTGTATATAAAGGTGCAAGTTCATTTTGAGGCTTTTCGATAA 1380  
DB 1490 AATTCAACCAAGATTTACTTGTATATAAAGGTGCAAGTTCATTTTGAGGCTTTTCGATAA 1549

QY 1381 CCGTTTGGAAACGACTCAAGGATGTCNCAAGTGTATAAGTCAAGTTAGTGGANGATTTCT 1440  
DB 1550 CCGTTTGGAAACGACTCAAGGATGTCNCAAGTGTATAAGTCAAGTTAGTGGANGATTTCT 1609

QY 1441 TCGCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTAGGAAAAACCAATTCGCAAAATTC 1500  
DB 1610 TCGCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTAGGAAAAACCAATTCGCAAAATTC 1669

QY 1501 CTACACTGATGATGAGATTCAGAGTAGCCAAAGTTTCGCGGCAAGTACCAACAGAGACGG 1560  
DB 1670 CTACACTGATGATGAGATTCAGAGTAGCCAAAGTTTCGCGGCAAGTACCAACAGAGACGG 1729

QY 1561 TTATATCTTTGATCTCTGATATTAACAGTGTAGAGGGGATGCTATGTAATCTCCACA 1620  
DB 1730 TTATATCTTTGATCTCTGATATTAACAGTGTAGAGGGGATGCTATGTAATCTCCACA 1789

QY 1621 TATGACCCATAGCCACTGGATTTAAAAAGATAGTTTGTCTGAAGCTGAGAGCGGCGAGC 1680  
DB 1790 TATGACCCATAGCCACTGGATTTAAAAAGATAGTTTGTCTGAAGCTGAGAGCGGCGAGC 1849

QY 1681 CCAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTTCGACAGACCATCAGGATTCAGG 1740  
DB 1850 CCAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTTCGACAGACCATCAGGATTCAGG 1909

QY 1741 AAATACTGAGCAAAAGGAGCAGAGCTATCTCAACCGCGTGAAGAGCAGCTAAGAGCT 1800  
DB 1910 AAATACTGAGCAAAAGGAGCAGAGCTATCTCAACCGCGTGAAGAGCAGCTAAGAGCT 1969

QY 1801 GCCACTTGTATGCTTTTCAATCTTCAATATCTGTAGAAGTCAAAAAACGGTAGTTT 1860  
DB 1970 GCCACTTGTATGCTTTTCAATCTTCAATATCTGTAGAAGTCAAAAAACGGTAGTTT 2029

1861	AATCATACCTCATTTATGACGATTAACCATTAACATCAAAATTTGAGTGGTTTGACGAAGCCCT	1920
1862		
1863		
1864		
1865		
1866		
1867		
1868		
1869		
1870		
1871		
1872		
1873		
1874		
1875		
1876		
1877		
1878		
1879		
1880		
1881		
1882		
1883		
1884		
1885		
1886		
1887		
1888		
1889		
1890		
1891		
1892		
1893		
1894		
1895		
1896		
1897		
1898		
1899		
1900		
1901		
1902		
1903		
1904		
1905		
1906		
1907		
1908		
1909		
1910		
1911		
1912		
1913		
1914		
1915		
1916		
1917		
1918		
1919		
1920		
1921	TTATGAGGCACCTAAGGGGTATACCTCTTGAGGATCTTTTGGCGACTCTCAAGTACTATGT	1980
1922		
1923		
1924		
1925		
1926		
1927		
1928		
1929		
1930		
1931		
1932		
1933		
1934		
1935		
1936		
1937		
1938		
1939		
1940		
1941		
1942		
1943		
1944		
1945		
1946		
1947		
1948		
1949		
1950		
1951		
1952		
1953		
1954		
1955		
1956		
1957		
1958		
1959		
1960		
1961		
1962		
1963		
1964		
1965		
1966		
1967		
1968		
1969		
1970		
1971		
1972		
1973		
1974		
1975		
1976		
1977		
1978		
1979		
1980		
1981	CGAATATCCAAACGAAACGCTCGCATTTCAGATAATGTTGGTAAACGCTAGCGACCATGT	2040
1982		
1983		
1984		
1985		
1986		
1987		
1988		
1989		
1990		
1991		
1992		
1993		
1994		
1995		
1996		
1997		
1998		
1999		
2000		
2001		
2002		
2003		
2004		
2005		
2006		
2007		
2008		
2009		
2010		
2011		
2012		
2013		
20		

SULT 7  
-10-412-862-5  
Sequence 5, Application US/10412862  
Publication No. US20040052781A1  
GENERAL INFORMATION:  
APPLICANT: Johnson, Leslie S.  
APPLICANT: Adamou, John E.  
TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
TITLE OF INVENTION: Motifs  
FILE REFERENCE: 469201-685  
CURRENT APPLICATION NUMBER: US/10/412,862  
CURRENT FILING DATE: 2003-04-14  
PRIOR APPLICATION NUMBER: 09/468,656  
PRIOR FILING DATE: 1999-12-21  
PRIOR APPLICATION NUMBER: 60/113,048  
PRIOR FILING DATE: 1998-12-21  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 3.0  
SEQ ID NO 5  
LENGTH: 2531  
TYPE: DNA  
ORGANISM: Streptococcus pneumoniae  
-10-412-862-5

Query Match	87.3%	Score 1999.6	DB 13	Length 2531	
Best Local Similarity	92.2%	Pred. No. 0			
Matches 2136	Conservative	0	Mismatches 154	Indels 27	Gaps 2
1	TTGTTCCTATGACCTTGGTCGTACCAAGCTGGTCAGGTTAAGAAAGTCTTAATCGAGT	60			
57	TTGTTCCTATGACCTTGGTCGTACCAAGCTGGTCAGGTTAAGAAAGTCTTAATCGAGT	116			
61	TTCTTATATAGATGGTGCATCAGGCTGGTCAAAAGGCCAGAAAACTTCACACCATCAAGT	120			
117	TTCTTATATAGATGGTGCATCAGGCTGGTCAAAAGGCCAGAAAACTTCACACCATCAAGT	176			
121	CAGTAAGAGGGAGGGGATCAACGCCGAAACAAATNGTNATCAAGATTCGGATCAAGTTA	180			
177	CAGTAAGAGGGAGGGGATCAACGCCGAAACAAATCGTCATCAAGATTACGGATCAAGTTA	236			

QY	181	TGTGACCTCTCATGGAGCACCATTATCAATTA	CTATAAATGGCGAAGGTTCTCTATGATGCCAT	244
Db	237	TGTGACCTCTCATGGAGCACCATTATCAATTA	CTATAAATGGCGAAGGTTCTCTATGATGCCAT	296
QY	241	CATCAGTGAAGAGCTCTCATGAAGAATCCGAA	TTATCAGTTGAGGATTCAGACATTTGT	300
Db	297	CATCAGTGAAGAGCTCTCATGAAGAATCCGAA	TTATCAGTTGAGGATTCAGACATTTGT	356
QY	301	CAATGAAATCAAGGGTGGTTATGTTCAATTA	AGGTTAAACGGTAAATCTATGTTATCTTAA	360
Db	357	CAATGAAATCAAGGGTGGTTATGTTCAAGTTA	TCAGTTAGATGGAATAATCTATGTTTAA	416
QY	361	GGATGAGCTCATCGGATTAATAATTCGACAAA	AGAGATTAAACGTCAAGACAGGA	420
Db	417	GGATGAGCTCATCGGATTAATAATTCGACAAA	AGAGATTAAACGTCAAGACAGGA	476
QY	421	ACGCAAGTCATAATCATAACTCAAGAGC---	AGATAATGCTGTTGCTGCAGCCAGAGCCCA	477
Db	477	ACACAGTCATAATCACGGGGGTGGTTCTTA	ACGATCAAGCAGTAGTTGCAGCCAGAGCCCA	536
QY	478	AGGACGTTATACAAACGATGATGGGTATATCT	TCCAATGCATCTCATATCATTTGAGGACAC	537
Db	537	AGGACGTTATACAAACGATGATGGGTATATCT	TCCAATGCATCTCATATCATTTGAGGACAC	596
QY	538	GGGTGATGCTTATATCGTTCTCTACGCGACCA	TTACCATTTACATTTCTTAAGAATGAGTT	597
Db	597	GGGTGATGCTTATATCGTTCTCTACGCGACCA	TTACCATTTACATTTCTTAAGAATGAGTT	656
QY	598	ATCAGCTTAGCGAGTTAGCTGTCAGAAAGCCT	TATGGAAATGGGAAGCAGGATCTCGTCC	657
Db	657	ATCAGCTTAGCGAGTTAGCTGTCAGAAAGCCT	TATGGAAATGGGAAGCAGGATCTCGTCC	716
QY	658	TTCTTCAAGTTCTAGTTTAAATGCAAAATCAG	ACTCAACCAAGATTGTCAGAGAACCACAA	717
Db	717	TTCTTCAAGTTCTAGTTTAAATGCAAAATCAG	ACTCAACCAAGATTGTCAGAGAACCACAA	776
QY	718	TCTGACTGTCACTCCAACTTATCATCAAAATC	AAAGGGAACCAATTTCAAGCCTTTTACG	777
Db	777	TCTGACTGTCACTCCAACTTATCATCAAAATC	AAAGGGAACCAATTTCAAGCCTTTTACG	836
QY	778	TGAATTTGATGCTTAAACCTTTATCAGAAACG	CGATGTGGAATCTGATGGCCTTATTTTCGA	837
Db	837	TGAATTTGATGCTTAAACCTTTATCAGAAACG	CGATGTGGAATCTGATGGCCTTATTTTCGA	896
QY	838	CCACGCGCAATCACAACTGCGAACCGCCAGAG	TGTAGCTGTCCTCATGTAACCAATTA	897
Db	897	CCACGCGCAATCACAACTGCGAACCGCCAGAG	TGTAGCTGTCCTCATGTAACCAATTA	956
QY	898	CCACTTTATCCCTTATGAAACAAATGTCTGA	ATTTGGAAAAACGAATTCCTGTTATTTTCC	957
Db	957	CCACTTTATCCCTTATGAAACAAATGTCTGA	ATTTGGAAAAACGAATTCCTGTTATTTTCC	1016
QY	958	CCTTGCTTATCGTTCAAAACATTTGGGTATCA	AGATTTAAGAACCGAACCAAGTCCACA	1017
Db	1017	CCTTGCTTATCGTTCAAAACATTTGGGTATCA	AGATTTAAGAACCGAACCGAACGATTCACA	1076
QY	1018	ATCGACTCCGGACCTAGTCCAAGTCGGCAACT	GCACCAAACTCTCAACGAGCTCCAG	1077
Db	1077	ATCGACTCCGGACCTAGTCCAAGTCGGCAACT	GCACCAAACTCTCAACGAGCTCCAG	1136
QY	1078	CAATCCAAATGATGAGAAATTTGTTCAAAGAG	CTGTTTCGAAAAAGTAGTGGCGATGTTATGT	1137
Db	1137	CAATCCAAATGATGAGAAATTTGTTCAAAGAG	CTGTTTCGAAAAAGTAGTGGCGATGTTATGT	1196
QY	1138	CTTTTGAGGAAATGGAGTTTCTCGTTTATATC	CCAGCCCAAGGATCTTTTCAGCAGAAAACG	1197
Db	1197	CTTTTGAGGAAATGGAGTTTCTCGTTTATATC	CCAGCCCAAGGATCTTTTCAGCAGAAAACG	1256
QY	1198	AGCAGCATTTGATGACAACTGGCCCAAGCAGAA	AGTTTATCTCATTAAGCTAGAGCTAA	1257
Db	1257	AGCAGCATTTGATGACAACTGGCCCAAGCAGAA	AGTTTATCTCATTAAGCTAGAGCTAA	1316

1258 GAAACTGACCTCCCATCTAGTGATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 1317  
1317 GAAACTGACCTCCCATCTAGTGATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 1376  
1318 AAGAAATTCACCAAGATTTACTTGATAATAAGGTCGACAGTTGATTTTGGCTTTGGA 1377  
1377 AAGAAATTCACCAAGATTTACTTGATAATAAGGTCGACAGTTGATTTTGGCTTTGGA 1436  
1378 TAACCTGTTGGAAACGACTCAAGGATGTCNCAAGTGTATAAAGTCAAGTTAGTGAGGATAT 1437  
1437 TAACCTGTTGGAAACGACTCAAGGATGTCNCAAGTGTATAAAGTCAAGTTAGTGAGGATAT 1496  
1438 TCTTGCCCTTTAGCTCCGATTCGTCATCCAGAAAGCTTTAGGAAACCAAAATCGCAAT 1497  
1497 TCTTGCCCTTTAGCTCCGATTCGTCATCCAGAAAGCTTTAGGAAACCAAAATCGCAAT 1556  
1498 TACCTACACTGATGATGAGATTCAGATAGCCAGTTGGCAGGCAAGTACACACAGAGA 1557  
1557 TACCTACACTGATGATGAGATTCAGATAGCCAGTTGGCAGGCAAGTACACACAGAGA 1616  
1558 CGGTTATATCTTGTATCTCGTGATATAACCAAGTGTAGAGGGGATGCCCTATGTAACCTC 1617  
1617 CGGTTATATCTTGTATCTCGTGATATAACCAAGTGTAGAGGGGATGCCCTATGTAACCTC 1676  
1618 ACATATGACCCATAGCCACTGGATTTAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGC 1677  
1677 ACATATGACCCATAGCCACTGGATTTAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGC 1736  
1678 AGCCAGGCTTATGCTTAAGAGAAAGTTTGACCCCTCTCTGACAGACATCAGGATTC 1737  
1737 AGCCAGGCTTATGCTTAAGAGAAAGTTTGACCCCTCTCTGACAGACATCAGGATTC 1796  
1738 AGGAAATCTGAGGCAAGAGGAGAGAGCTATCTACAAACCGCTGAAAGAGCTTAAGAA 1797  
1797 AGGAAATCTGAGGCAAGAGGAGAGAGCTATCTACAAACCGCTGAAAGAGCTTAAGAA 1856  
1798 GGTGCACCTTGATCGTATGCTTCAATCTTCAATATATCTGTAAGAGTCAAAACCGGTAG 1857  
1857 GGTGCACCTTGATCGTATGCTTCAATCTTCAATATATCTGTAAGAGTCAAAACCGGTAG 1916  
1858 TTTAATCATCTCTAATATGACCATTAACATAACATCAAAATTTGAGTGGTTTGAACGAGG 1917  
1917 TTTAATCATCTCTAATATGACCATTAACATAACATCAAAATTTGAGTGGTTTGAACGAGG 1976  
1918 CCTTTATGAGGACCTTAAGGCTTACTCTGAGGATCTTTTGGCGACTGTCAAGTACTA 1977  
1977 CCTTTATGAGGACCTTAAGGCTTACTCTGAGGATCTTTTGGCGACTGTCAAGTACTA 2036  
1978 TGTGAAACATCCAAACGAAACGTCGCAATTCAGATATGTTTGGTAAACGCTAGCGACCA 2037  
2037 TGTGAAACATCCAAACGAAACGTCGCAATTCAGATATGTTTGGTAAACGCTAGCGACCA 2096  
2038 TGTTCAAAGAAACAAAAATGTCAGCTGATACCAATCAAAACGGAATAACCAAGCGAGGA 2097  
2097 TGTTCGTAATAAAGGTAGACCAAGACAGTAAACCTGATGAAGATAAGGAAACATGATGA 2156  
2098 GAAACCTCAGACAGAAACCTGAGGAAGA-----AACCC 2133  
2157 AGTAAGTGAAGCAACTCAGCTTGAATCTGTAAGAAAGAGAAATCAGCTGGTTTAAATCC 2216  
2134 TCGAGAGAGAAACCGAAGCGAGAAACAGAGTCTCCAAACCAACCAAGAGAGAAACGAGA 2193  
2217 TTCAGAGATAATCTTTTAAACCAAGACTGATACCGAAGAGACAGAGAGAGAGAGCTGA 2276  
2194 AGAATCACCAAGGAATCAGAGAAACCTCAGGTCGAGACTGAAAAAGGTTTCAAGAAAAACT 2253  
2277 AGATACACAGATGAGGCTGAATCTCTCAAGTAGAGAAATCTGTTTATTAACCGTAAAGT 2336  
2254 GAGAGAGGCTGAAGATTTACTTTGGAAAAATCCAGGAT 2290  
2337 AGCAGATGCGGAGGCTTCTGTAGAAAAAGTAAACAGAT 2373

RESULT 8  
US-10-412-850-5  
; Sequence 5, Application US/10412850  
; Publication No. US2004001836A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; TITLE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-686  
; CURRENT APPLICATION NUMBER: US/10/412,850  
; CURRENT FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 09/468,656  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 5  
; LENGTH: 2531  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-10-412-850-5  
Query Match 87.3%; Score 1999.6; DB 16; Length 2531;  
Best Local Similarity 92.2%; Pred. No. 0;  
Matches 2136; Conservative 0; Mismatches 154; Indels 27; Gaps 2;  
QY 1 TTGTTCTATGAATCGTCTGTCACCAAGCTGGTCAAGTAAAGAGAGTCTTAATCGAGT 60  
Db 57 TTGTTCTATGAATCGTCTGTCACCAAGCTGGTCAAGTAAAGAGAGTCTTAATCGAGT 116  
QY 61 TTCTTATATAGATGGTGATCGGCTGTCACCAAGCTGGTCAAGTAAAGAGAGTCTTAATCGAGT 120  
Db 117 TTCTTATATAGATGGTGATCGGCTGTCACCAAGCTGGTCAAGTAAAGAGAGTCTTAATCGAGT 176  
QY 121 CAGTAAGAGAGGAGGATCAACCGGCAAGTCAAGTAAAGAGTCAAGTAAAGAGTCAAGTAA 180  
Db 177 CAGTAAGAGAGGAGGATCAACCGGCAAGTCAAGTAAAGAGTCAAGTAAAGAGTCAAGTAA 236  
QY 181 TGTGACCTCTCATGAGACCATTCATTAATAATGCGCAAGGTTCCCTTATGATGCCAT 240  
Db 237 TGTGACCTCTCATGAGACCATTCATTAATAATGCGCAAGGTTCCCTTATGATGCCAT 296  
QY 241 CATCAGTGAAGAGCTCCTCATGAAGATCCGAAATTCAGTTGAAGATTCAGACATGTT 300  
Db 297 CATCAGTGAAGAGCTCCTCATGAAGATCCGAAATTCAGTTGAAGATTCAGACATGTT 356  
QY 301 CAATGAATCAAGGCTGTTATGTCATTAAGTAAACGTTAAATCTATGTTTACCTTAA 360  
Db 357 CAATGAATCAAGGCTGTTATGTCATTAAGTAAACGTTAAATCTATGTTTACCTTAA 416  
QY 361 GGATGACGCTCATGCGGATTAATTCGGAACAAAGAGAGATTAACCTCAGAACGAGA 420  
Db 417 GGATGACGCTCATGCGGATTAATTCGGAACAAAGAGAGATTAACCTCAGAACGAGA 476  
QY 421 ACGAGTCAATATCAATCAAGAGC---AGATAATGCTGTTGCTGAGCAGAGCCCA 477  
Db 477 ACGAGTCAATATCAAGAGC---AGATAATGCTGTTGCTGAGCAGAGCCCA 536  
QY 478 AGACGTTATCAACGAGTATGGTATATCTTCAATGCATCTGATATCATTTAGAGCAC 537  
Db 537 AGACGTTATCAACGAGTATGGTATATCTTCAATGCATCTGATATCATTTAGAGCAC 596  
QY 538 GGTGTGCTTATATCTTCTCAGCGGACCATTCACATTAATTCCTTAAGATGAGTT 597  
Db 597 GGTGTGCTTATATCTTCTCAGCGGACCATTCACATTAATTCCTTAAGATGAGTT 656  
QY 598 ATCAGCTAGCGAGTATGCTGTCAGAGGCTTATGGAATGGGAGCAGGAGTCTGTC 657  
Db 657 ATCAGCTAGCGAGTATGCTGTCAGAGGCTTATGGAATGGGAGCAGGAGTCTGTC 716



```
QY 658 TTCTTCAAGTTCTAGTTTATATGTAATCCAGCTCAACCAAGATTGTTCAGAGAACCAAA 717
Db 717 TTCTTCAAGTTCTAGTTTATATGTAATCCAGCTCAACCAAGATTGTTCAGAGAACCAAA 776.
QY 718 TCTGACTGTCACTCCCACTTATCATCAAAATCAAGGGGAAACATTTCAGCCCTTTTACG 777
Db 777 TCTGACTGTCACTCCCACTTATCATCAAAATCAAGGGGAAACATTTCAGCCCTTTTACG 836
QY 778 TGAATTGTATGCTAAACCCCTTATCAGAACGCCATGTGGAAATCTGATGGCCCTTATTTTGA 837
Db 837 TGAATTGTATGCTAAACCCCTTATCAGAACGCCATGTGGAAATCTGATGGCCCTTATTTTGA 896
QY 838 CCAGCGCAAAATCACAAGTGCAGACCGCGAGAGTGTAGCTGCTCCTCATGTGTAAACATTATTA 897
Db 897 CCAGCGCAAAATCACAAGTGCAGACCGCGAGAGTGTAGCTGCTCCTCATGTGTAAACATTATTA 956
QY 898 CCACCTTTATCCCTTATGAAAGATGCTGAATTCGAAAGAGAAATTCGCTGATTTATTTCC 957
Db 957 CCACCTTTATCCCTTATGAAAGATGCTGAATTCGAAAGAGAAATTCGCTGATTTATTTCC 1016
QY 958 CTTTTCGTTATCGTTCAAAACCATTCGGGTACCAAGATTCAGAGCAGAAACCAACCAAGTCCACA 1017
Db 1017 CTTTTCGTTATCGTTCAAAACCATTCGGGTACCAAGATTCAGAGCAGAAACCAACCAAGTCCACA 1076
QY 1018 ATCGACTCCGGAACCTAGTCCAGTCCGCAACCTGCACCAAAATCCTCAACAGCTCCAAAG 1077
Db 1077 ATCGACTCCGGAACCTAGTCCAGTCCGCAACCTGCACCAAAATCCTCAACAGCTCCAAAG 1136
QY 1078 CAATCCAAATGATGAGAAATTTGTTCAAGAGCTGTTTCGAAAGTAGGCGCATGTTATGT 1137
Db 1137 CAATCCAAATGATGAGAAATTTGTTCAAGAGCTGTTTCGAAAGTAGGCGCATGTTATGT 1196
QY 1138 CTTTGTAGGAAATGAGCTTCTCGTTATATCCAGCCAAAGATCTTTCAGCAGAAAAACAGC 1197
Db 1197 CTTTGTAGGAAATGAGCTTCTCGTTATATCCAGCCAAAGATCTTTCAGCAGAAAAACAGC 1256
QY 1198 AGCAGCAATGATAGCAAACTGGCCAAAGCAGGAAAGTTTATCTCATAGCTTAGGAGCTAA 1257
Db 1257 AGCAGCAATGATAGCAAACTGGCCAAAGCAGGAAAGTTTATCTCATAGCTTAGGAGCTAA 1316
QY 1258 GAAACCTGACTCCCATCTAGTGTATGAGAAATTTTACAAATAGGCTTATGATCTTAGC 1317
Db 1317 GAAACCTGACTCCCATCTAGTGTATGAGAAATTTTACAAATAGGCTTATGATCTTAGC 1376
QY 1318 AAGAAATTCACCAAGATTTACTTGATATAAAGGTCGACAAAGTTCATTTTGGAGCTTTTGA 1377
Db 1377 AAGAAATTCACCAAGATTTACTTGATATAAAGGTCGACAAAGTTCATTTTGGAGCTTTTGA 1436
QY 1378 TAACCTGTTGGAAACGACTCAAGGATGTCNCAAGTGATAAAGTCAAGTTAGTGGANGATAT 1437
Db 1437 TAACCTGTTGGAAACGACTCAAGGATGTCNCAAGTGATAAAGTCAAGTTAGTGGANGATAT 1496
QY 1438 TCTTGCCTTCTTAGCTCCGATTGCTCATCCAGACGTTTAGGAAACCAATGGCGCAAT 1497
Db 1497 TCTTGCCTTCTTAGCTCCGATTGCTCATCCAGACGTTTAGGAAACCAATGGCGCAAT 1556
QY 1498 TACCTACACTGATGATGAGATTCAAGTAGCCCAAGTTGGCGGCAAGTACACAAACAGAGA 1557
Db 1557 TACCTACACTGATGATGAGATTCAAGTAGCCCAAGTTGGCGGCAAGTACACAAACAGAGA 1616
QY 1558 CGGTTATATCTTTGATCCTCGTGATATAACCAAGTATGAGGGGATGCTATGTAACTCC 1617
Db 1617 CGGTTATATCTTTGATCCTCGTGATATAACCAAGTATGAGGGGATGCTATGTAACTCC 1676
QY 1618 ACATATGACCCATAGCCACTGGATTAAAAAGATAGTTTGTGAAGCTGAGAGCGGCG 1677
Db 1677 ACATATGACCCATAGCCACTGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAGCGGCG 1736
QY 1678 AGCCGAGGCTTATGCTAAAGAGAAAGTTTGAACCCCTCCTTCGACAGACCATCAGGATTC 1737
Db 1737 AGCCGAGGCTTATGCTAAAGAGAAAGTTTGAACCCCTCCTTCGACAGACCATCAGGATTC 1796
QY 1738 AGGAAATACTGAGGCAAAAGAGCAGAGCTATCTACAAACCGCGTGAAGCAGCTAAGAA 1797
```

```
Db 1797 AGGAAATACTGAGGCAAAAGAGCAGAGCTATCTACAAACCGGTGAAACGCTAAGAA 1856
QY 1798 GGTGCACTTGTATGCTATGCTTACAAATCTTCAATATATCTGTAGAAAGTCAAAAACGGTAG 1857
Db 1857 GGTGCACTTGTATGCTATGCTTACAAATCTTCAATATATCTGTAGAAAGTCAAAAACGGTAG 1916
QY 1858 TTTAATCATACCTCATATATGACCATTTACCAATCAATCAATCAATTTGAGTGGTTTCAAGAGG 1917
Db 1917 TTTAATCATACCTCATATATGACCATTTACCAATCAATCAATTTGAGTGGTTTCAAGAGG 1976
QY 1918 CTTTATGAGGCACTTAAGGGGTATCTCTTGAAGATCTTTTGGCGACTGTCTCAAGTACTA 1977
Db 1977 CTTTATGAGGCACTTAAGGGGTATCTCTTGAAGATCTTTTGGCGACTGTCTCAAGTACTA 2036
QY 1978 TGTCCAAACATCCAAACGAAAGCTCCGCAATTCAGATTCAGATATGTTTTTGGTAAACGCTAGCGACCA 2037
Db 2037 TGTCCAAACATCCAAACGAAAGCTCCGCAATTCAGATTCAGATATGTTTTTGGTAAACGCTAGCGACCA 2096
QY 2038 TGTTCAAAGAAACAAAAATGTTCAAGCTGATACCAATCAAAACGAAACCAAGCGAGGA 2097
Db 2097 TGTTCGTAATAATAAGGTAGACCAAGACAGTAAACCTGATGAAGATAAGGAACATGATGA 2156
QY 2098 GAAACCTCAGACAGAAAAACCTTGAGGAAGA-----AACCCC 2133
Db 2157 AGTAAAGTGAGCCAACTCAACCTGGAATCTGATGAAAAAGAGAATCAACGCTGGTTTAAATCC 2216
QY 2134 TCGAGAAGAGAAACCGCAAGCGAGAAACAGAGTCTCCAAACCAACACAGAGAACCGAGA 2193
Db 2217 TTCAGCAGATATCTTTTATAAACCAAGCACTGATACGGAAGACAGAGAGGAAGAAGCTGA 2276
QY 2194 AGAATCACAGAGGAATCAGAAAGAACCTCAGGTCGAGACTGAAAAAGGTTGAAGAAAAACT 2253
Db 2277 AGATACCAACAGATGAGGCTGAAATTCCTCAAGTAGAGAATTCGTTTATTAACGCTAAGAT 2336
QY 2254 GAGAGAGGCTGAAGATTTACTTTGGAAAAATCCAGGAT 2290
Db 2337 AGCAGATCGGAGGCGCTTGCTAGAAAAAGTAACAGAT 2373
```

## RESULT 9

```
US-10-387-783-5
; Sequence 5, Application US/10387783
; Publication No. US2004000531A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-683
; CURRENT APPLICATION NUMBER: US/10/387,783
; CURRENT FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-387-783-5
```

Query Match 87.3%; Score 1999.6; DB 16; Length 2531;  
Best Local Similarity 92.2%; Pred. No. 0;  
Matches 2136; Conservative 0; Mismatches 154; Indels 27; Gaps 2;

```
QY 1 TTGTTCTTATGAACCTTGGTCTGTCACCAAGCTGGTCAGGTTAAGAAAGTCTTAATCAGT 60
Db 57 TTGTTCTTATGAACCTTGGTCTGTCACCAAGCTGGTCAGGTTAAGAAAGTCTTAATCAGT 116
```

QY 61 TTCTTATATAGATGGTGTGATCAGGCTGGTCAAAAGCGCAGAAACCTTTGACACACAGATGAAGT 120  
Db 117 TTCTTATATAGATGGTGTGATCAGGCTGGTCAAAAGCGCAGAAACCTTTGACACACAGATGAAGT 176  
QY 121 CAGTAAGAGGAGGAGGATCAACGCGCAACAATGWTNATCAAGATTACGGATCAAGGTTA 180  
Db 177 CAGTAAGAGGAGGAGGATCAACGCGCAACAATGWTNATCAAGATTACGGATCAAGGTTA 236  
QY 181 TGTGACCTCTCATGAGACCATTAATCAATTAATAATGGCAAGGTTCCCTTATGATGCCAT 240  
Db 237 TGTGACCTCTCATGAGACCATTAATCAATTAATAATGGCAAGGTTCCCTTATGATGCCAT 296  
QY 241 CATCAGTGAAGAGCTCCTCATGAAGAGTCCGAATATCAGTTGAGAGATTCAGACATTTG 300  
Db 297 CATCAGTGAAGAGCTCCTCATGAAGAGTCCGAATATCAGTTGAGAGATTCAGACATTTG 356  
QY 301 CAATGAATCAAGGCTGTTATGTCATTAAGTAAACGGTAAATACATATGTTACCTTAA 360  
Db 357 CAATGAATCAAGGCTGTTATGTCATTAAGTAAACGGTAAATACATATGTTACCTTAA 416  
QY 361 GATGACAGCTCATGCGGATTAATATTCGGAACAAAAGAGAGATTAAACGTCAGAGCAGGA 420  
Db 417 GGATGACGCTCATGCGGATTAATATTCGGAACAAAAGAGAGATTAAACGTCAGAGCAGGA 476  
QY 421 ACGAGTCATTAATCAATCAAGAGC---AGATPAATGCTGTTGCTGACGACGAGGCCA 477  
Db 477 ACACAGTCATAATCAACGGGGGTGTTCTAACGATCAAGCAGTAGTTGCGACGACAGGCCA 536  
QY 478 AGGACGTTATACACAGATGATGGTATATCTTCAATGCATCTCATATCATTTGAGGACAC 537  
Db 537 AGGACGTTATACACAGATGATGGTATATCTTCAATGCATCTCATATCATTTGAGGACAC 596  
QY 538 GGTGATGCTTATATCTGCTCACGCGGACCATTTACCATTTACATTTCTTAAGATGAGTT 597  
Db 597 GGTGATGCTTATATCTGCTCACGCGGACCATTTACCATTTACATTTCTTAAGATGAGTT 656  
QY 598 ATCAGTAGCGATGATGCTGCTCAGAGAGCCTATTTGGAATGGGAAGCAGGATCTCGTCC 657  
Db 657 ATCAGTAGCGATGATGCTGCTCAGAGAGCCTATTTGGAATGGGAAGCAGGATCTCGTCC 716  
QY 658 TTCTTCAAGTTCTAGTTATTAATGGAATCCAGCTCAACCGAGATTGTCAGAGAACCCAA 717  
Db 717 TTCTTCAAGTTCTAGTTATTAATGGAATCCAGCTCAACCGAGATTGTCAGAGAACCCAA 776  
QY 718 TCTGACTGTCACCTCCAACTTATCATCAAAATCAGGGGAAACATTTCAAGCCTTTACG 777  
Db 777 TCTGACTGTCACCTCCAACTTATCATCAAAATCAGGGGAAACATTTCAAGCCTTTACG 836  
QY 778 TGAATTTGATGCTTAAACCTTTATCAGAACGCCATGTGGAATCTGATGGCCTTATTTTGA 837  
Db 837 TGAATTTGATGCTTAAACCTTTATCAGAACGCCATGTGGAATCTGATGGCCTTATTTTGA 896  
QY 838 CCCAGGCGAATCACAAGTCGAACCGGCGAGAGTGTAGCTGTCCCTCATGGTAAACCATTA 897  
Db 897 CCCAGGCGAATCACAAGTCGAACCGGCGAGAGTGTAGCTGTCCCTCATGGTAAACCATTA 956  
QY 898 CCACTTTATCCCTTATGAAACAAATGCTGAATGGAATAAGGAAACGATTTGCTCGTATTATCC 957  
Db 957 CCACTTTATCCCTTATGAAACAAATGCTGAATGGAATAAGGAAACGATTTGCTCGTATTATCC 1016  
QY 958 CTTTCTGTTATCGTTCAAAACCATTTGAGTACAGATTTCAAGACCCAGAACCAACCAAGTCCACA 1017  
Db 1017 CTTTCTGTTATCGTTCAAAACCATTTGAGTACAGATTTCAAGACCCAGAACCAACCAAGTCCACA 1076  
QY 1018 ATCGACTCCGGAACCTTAGTCCAAAGTCGCAACCTGACACCAAAATCTTCAACCGAGTCCAAAG 1077  
Db 1077 ATCGACTCCGGAACCTTAGTCCAAAGTCGCAACCTGACACCAAAATCTTCAACCGAGTCCAAAG 1136  
QY 1078 CAAATCAATTTGATGAAATTTGTTCAAGAGCTGTTTCAAAAAGTAGGCGATGTTATGT 1137  
Db 1137 CAAATCAATTTGATGAAATTTGTTCAAGAGCTGTTTCAAAAAGTAGGCGATGTTATGT 1196  
QY 1138 CTTTGGAGGAGATGGAGTTTCTCGTTATATATCCAGCGCAGGATCTTTTCAGCAGAACACG 1197

Db 1197 CTTTGGAGGAGATGGAGTTTCTCGTTATATATCCAGCGCAAGATCTTTTCAGCAGAACACG 1256  
QY 1198 AGCAGGCAATGATAGCAAACTGCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAA 1257  
Db 1257 AGCAGGCAATGATAGCAAACTGCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAA 1316  
QY 1258 GAAACCTGACCTCCCATCTAGTATCGAGAAATTTTCAATAAGGCTTATGACTTACTAGC 1317  
Db 1317 GAAACCTGACCTCCCATCTAGTATCGAGAAATTTTCAATAAGGCTTATGACTTACTAGC 1376  
QY 1318 AAGAAATCCCAAGATTTTACTTGTATTAATAAGGTTCGACAAAGTTGATTTTTCAGGCTTTGA 1377  
Db 1377 AAGAAATCCCAAGATTTTACTTGTATTAATAAGGTTCGACAAAGTTGATTTTTCAGGCTTTGA 1436  
QY 1378 TAACTCTGTTGGAACGACTCAAGGATGTCNCAAGTGTAAAGTCAAGTTAGTGGAGATAT 1437  
Db 1437 TAACTCTGTTGGAACGACTCAAGGATGTCNCAAGTGTAAAGTCAAGTTAGTGGAGATAT 1496  
QY 1438 TCTTGGCTTCTTTAGCTCCGATTCGATCCAGAACTTTTAGGAAACCAATCGCAAT 1497  
Db 1497 TCTTGGCTTCTTTAGCTCCGATTCGATCCAGAACTTTTAGGAAACCAATCGCAAT 1556  
QY 1498 TACCTACACTGATGATGAGATTCAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAGA 1557  
Db 1557 TACCTACACTGATGATGAGATTCAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAGA 1616  
QY 1558 CGGTTATATCTTTGATCTCTGATATAACCAAGTATGAGGGGATGCTTATGTAATCTCC 1617  
Db 1617 CGGTTATATCTTTGATCTCTGATATAACCAAGTATGAGGGGATGCTTATGTAATCTCC 1676  
QY 1618 ACATATGACCCATGACCCACTGGATTAATAAGATAGTTTGTCTGAAGCTGAGAGCGGC 1677  
Db 1677 ACATATGACCCATGACCCACTGGATTAATAAGATAGTTTGTCTGAAGCTGAGAGCGGC 1736  
QY 1678 AGCCAGGCTTATGCTTAAAGAGAAAGTTTGAACCTCTTCTGACAGACCATCAGGATTC 1737  
Db 1737 AGCCAGGCTTATGCTTAAAGAGAAAGTTTGAACCTCTTCTGACAGACCATCAGGATTC 1796  
QY 1738 AGGAAATCTAGGCAAAAGGAGCAGAGCTATCTACACCGCTGAAAGCAGCTAAGAA 1797  
Db 1797 AGGAAATCTAGGCAAAAGGAGCAGAGCTATCTACACCGCTGAAAGCAGCTAAGAA 1856  
QY 1798 GGTGCACTTGTATGCTTACATCTTCAATATATCTGTAGAAAGTCAAAAACGCTAG 1857  
Db 1857 GGTGCACTTGTATGCTTACATCTTCAATATATCTGTAGAAAGTCAAAAACGCTAG 1916  
QY 1858 TTTTATCATCTCATTTATGACCATTTACCATTAACATTTTGGCTGTTTTCAGCAGG 1917  
Db 1917 TTTTATCATCTCATTTATGACCATTTACCATTAACATTTTGGCTGTTTTCAGCAGG 1976  
QY 1918 CTTTATGAGCACCCTAAGGGGTATCTCTTGGAGTCTTTTGGCGACTGTCAAGTACTA 1977  
Db 1977 CTTTATGAGCACCCTAAGGGGTATCTCTTGGAGTCTTTTGGCGACTGTCAAGTACTA 2036  
QY 1978 TGTGAAATCTCAAAAGAAAGCTCCGATTTAGATTAATGTTTTTGGTAAACGCTAGCAGCA 2037  
Db 2037 TGTGAAATCTCAAAAGAAAGCTCCGATTTAGATTAATGTTTTTGGTAAACGCTAGCAGCA 2096  
QY 2038 TGTTCAGAAACCAAAATGTTGATGATACCAATCAACCGGAAACCAACCAAGCAGGA 2097  
Db 2097 TGTTCAGAAACCAAAATGTTGATGATACCAATCAACCGGAAACCAACCAAGCAGGA 2156  
QY 2098 GAAACCTCAGACAGAAACCTGAGGAAAG-----AACCCC 2133  
Db 2157 AGTAAGTGAACCACTCACCTCTGATCTGTAAGAAAGAGATCAGCTGGTTTAAATCC 2216  
QY 2134 TCGAGAGAGAGAAACCGCAAAAGCAGGAAACCAAGTCTTCAAAACCAACAGAGAAACCGA 2193  
Db 2217 TTCAGCAGATTAATCTTTATATAACCAAGCACTGATACGGAAGAGACAGAGAAAGAGCTGA 2276  
QY 2194 AGATCACCAGAGAAATCAGAAAGAACCTCAGGTCCAGACTGAAAAGGTTGAAAGAAAACCT 2253

Db 2277 AGATACACAGATGAGGCTGAATTCCTCAAGTAGAGAAATTCCTGTTATTAACGCTAAGAT 2336  
QY 2254 GAGAGAGCTGAAGATTACTTTGGAATAATCCAGGAT 2290  
Db 2337 AGCAGATGCGAGGCTTGCTAGAAAAGTAAACAGAT 2373  
RESULT 10  
US-10-158-844-94  
; Sequence 94, Application US/10158844  
; Publication No. US20040029118A1  
; GENERAL INFORMATION:  
; APPLICANT: Kunsch et al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-R  
; COMPUTER: Dell Latitude Pentium 3  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/158,844  
; FILING DATE: 03-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/961,527  
; FILING DATE: 1997-10-30  
; APPLICATION NUMBER: US 60/029,960  
; FILING DATE: 1996-10-31  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hyman, Mark J.  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB340P1D1  
; INFORMATION FOR SEQ ID NO: 94:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8195 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 94:  
US-10-158-844-94  
Query Match 87.3%; Score 1999.2; DB 13; Length 8195;  
Best Local Similarity 92.1%; Pred. No. 0;  
Matches 2135; Conservative 1; Mismatches 154; Indels 27; Gaps 2;  
QY 1 TTGTTCTCTAAGAACTTGCTGCTCAAGCTGGTCAAGGTTAAGAAAGAGTCTAATCGAGT 60  
Db 5658 TTGTTCTCTAAGAACTTGCTGCTCAAGCTGGTCAAGGTTAAGAAAGAGTCTAATCGAGT 5717  
QY 61 TTCTTATATAGATGGTATCAGGCTGGTCAAGGTTAAGAAAGAGTCTAATCGAGT 120  
Db 5718 TTCTTATATAGATGGTATCAGGCTGGTCAAGGTTAAGAAAGAGTCTAATCGAGT 5777  
QY 121 CAGTAAGAGGAGGAGGATCAAGCGGCAAAATGCTATCAAGATTAAGGTTAAGGTTA 180  
Db 5778 CAGTAAGAGGAGGAGGATCAAGCGGCAAAATGCTATCAAGATTAAGGTTAAGGTTA 5837  
QY 181 TGTGACCTCTCATGAGACCAATTATCACTATTAATGGCAAGGTTCTCTTATGATGCCAT 240  
Db 5838 TGTGACCTCTCATGAGACCAATTATCACTATTAATGGCAAGGTTCTCTTATGATGCCAT 5897  
QY 241 CATCAGTGAAGAGTCTCTCATGAAGAGTCCGAATTAATCAGTTGAAGAGTTCAGACATTTG 300  
Db 5898 CATCAGTGAAGAGTCTCTCATGAAGAGTCCGAATTAATCAGTTGAAGAGTTCAGACATTTG 5957

QY 301 CAATGAAATCAAGGGTGGTTATGTCATTAAGGTAAACCGGTAAATATCTATCTGTTACCTTAA 360  
Db 5958 CAATGAAATCAAGGGTGGTTATGTCATTAAGGTAAACCGGTAAATATCTATCTGTTACCTTAA 6017  
QY 361 GGATGAGCTCATGCGGATAATATTCGGACAAAAGAGAGATTAAACGTCAGAAAGCAGGA 420  
Db 6018 GGATGAGCTCATGCGGATAATATTCGGACAAAAGAGAGATTAAACGTCAGAAAGCAGGA 6077  
QY 421 ACGCAGTCATATCAATCAATCAAGAGC--AGATAATGCTGTTGCTGCAGCCAGAGCCCA 477  
Db 6078 ACACAGTCATATCAAGGGGGTGGTTCTTAACGATCAAGCAGTAGTTGCGACCGAGCCCA 6137  
QY 478 AGGACGTTATCAACAGGATGATGGGTATATCTTCAATGCAATCTGATATCAATGAGGACAC 537  
Db 6138 AGGACGTTATCAACAGGATGATGGTTATATCTTCAATGCAATCTGATATCAATGAGGACAC 6197  
QY 538 GGGTATGCTTATATGCTTCTCACGGGACCAATTACCATTAATTCCTTAAGAAATGAGTT 597  
Db 6198 GGGTATGCTTATATGCTTCTCACGGGACCAATTACCATTAATTCCTTAAGAAATGAGTT 6257  
QY 598 ATCAGCTAGCGAGTTAGCTGCTGCAGAGCCTATTGGAATGGGAAGCAGGATCTCTGCC 657  
Db 6258 ATCAGCTAGCGAGTTAGCTGCTGCAGAGCCTATTGGAATGGGAAGCAGGATCTCTGCC 6317  
QY 658 TTCTTCAAGTTCTAGTTATAATGCAAAATCCAGCTCAACCAAGATTGTGACAGAACCAAA 717  
Db 6318 TTCTTCAAGTTCTAGTTATAATGCAAAATCCAGCTCAACCAAGATTGTGACAGAACCAAA 6377  
QY 718 TCTGACTGTCACCTCCAACTTATCATCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACG 777  
Db 6378 TCTGACTGTCACCTCCAACTTATCATCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACG 6437  
QY 778 TCAATTTGATGCTTAACCTTATCAGAGCCCATGTTGGAATCTGATGGCCCTTATTTTGA 837  
Db 6438 TGAATTTGATGCTTAACCTTATCAGAGCCCATGTTGGAATCTGATGGCCCTTATTTTGA 6497  
QY 838 CCCAGCGCAAAATCAAGTCCGAACCCAGAGGTGTAGCTGTCCCTCATGTTAACCATT 897  
Db 6498 CCCAGCGCAAAATCAAGTCCGAACCCAGAGGTGTAGCTGTCCCTCATGTTAACCATT 6557  
QY 898 CCACCTTTATCCCTTATGAAACAAATGTCGAAATTTGGAACAAATGCTGCTATTTATTC 957  
Db 6558 CCACCTTTATCCCTTATGAAACAAATGTCGAAATTTGGAACAAATGCTGCTATTTATTC 6617  
QY 958 CCTTCGTTATCGTTCAACCATTTGGTACAGATTCAGACCCAGAACCAACCAAGTCCACA 1017  
Db 6618 CCTTCGTTATCGTTCAACCATTTGGTACAGATTCAGACCCAGAACCAACCAAGTCCACA 6677  
QY 1018 ATCGACTCCGGAACCTTAGTCCGAACCTCGCAACCTCGCAACCTCCTCAACCCAGCTCCAAG 1077  
Db 6678 ATCGACTCCGGAACCTTAGTCCGAACCTCGCAACCTCGCAACCTCCTCAACCCAGCTCCAAG 6737  
QY 1078 CAATCCAATTGATGAGAAATTTGTTCAAGAAAGCTGTTTCGAAAAGTAGGCGATGTTATGT 1137  
Db 6738 CAATCCAATTGATGAGAAATTTGTTCAAGAAAGCTGTTTCGAAAAGTAGGCGATGTTATGT 6797  
QY 1138 CTTTGAAGGAGATGAGTTTCTCGTTATATATCCAGCCAGGATCTTTTCAGCAGAAACAGC 1197  
Db 6798 CTTTGAAGGAGATGAGTTTCTCGTTATATATCCAGCCAGGATCTTTTCAGCAGAAACAGC 6857  
QY 1198 AGCAGGATTTGATAGCAAACTGSCCAAGCAGGAAAGTTTATCTCATAGCTTAGGAGCTAA 1257  
Db 6858 AGCAGGATTTGATAGCAAACTGSCCAAGCAGGAAAGTTTATCTCATAGCTTAGGAGCTAA 6917  
QY 1258 GAAAACCTGACCTCCCATCTAGTGATCGAGAAATTTTCAATAAGGCTTATGACTTACTAGC 1317  
Db 6918 GAAAACCTGACCTCCCATCTAGTGATCGAGAAATTTTCAATAAGGCTTATGACTTACTAGC 6977  
QY 1318 AAGAAATTCACCAAGATTACTGTTAATTAAGGTCGACAGTTGATTTTGAAGGCTTTTGA 1377  
Db 6978 AAGAAATTCACCAAGATTACTGTTAATTAAGGTCGACAGTTGATTTTGAAGGCTTTTGA 7037  
QY 1378 TAACCTGTTGGAACGACTCAAGGATGTCNCAAGTGTAAAGTCAAGTCTAGTGGANGATAT 1437

7038	TAACTGTTGAACGACTCAAGGATGTCTCAAGTGATTAAGTCAAGTTAGTGGATGATAT	7097
1438	TCCTTGCCTTCTTAGCTCCGATTCGATCCAGAAAGTTTAGGAAACCAATATGCGCAAT	1497
7098	TCCTTGCCTTCTTAGCTCCGATTCGATCCAGAAAGTTTAGGAAACCAATATGCGCAAT	7157
1498	TACCTACACTGATGATGAGATTCAAAGTGCACAAAGTTGGCAGGCAAGTACAAACAGAAGA	1557
7158	TACCTACACTGATGATGAGATTCAAAGTGCACAAAGTTGGCAGGCAAGTACAAACAGAAGA	7217
1558	CGGTATATCTTTGATCCTCGTGATATACCAAGTATGAGGGGGATGCCATGTAACTTC	1617
7218	CGGTATATCTTTGATCCTCGTGATATACCAAGTATGAGGGGGATGCCATGTAACTTC	7277
1618	ACATATGACCCATAGCCACATGGGATTAATAAAGATAGTTTGTCTGAAGCTGAGAGCGCGC	1677
7278	ACATATGACCCATAGCCACATGGGATTAATAAAGATAGTTTGTCTGAAGCTGAGAGCGCGC	7337
1678	AGCCCGAGCTTATGCTAAAGAGAAAGTTTGACCCCTCTCTCGACAGACCATCAGGATTC	1737
7338	AGCCCGAGCTTATGCTAAAGAGAAAGTTTGACCCCTCTCTCGACAGACCATCAGGATTC	7397
1738	AGGAAATACTGAGCAAAAGGAGCAAGACTATCTACAAACCGCTGAAAGCAGCTAAGAA	1797
7398	AGGAAATACTGAGCAAAAGGAGCAAGACTATCTACAAACCGCTGAAAGCAGCTAAGAA	7457
1798	GGTGGCACTTGATCGTATGCCCTTACCAATCTTCAATATACTGTAGAGTCAAAACCGGTAG	1857
7458	GGTGGCACTTGATCGTATGCCCTTACCAATCTTCAATATACTGTAGAGTCAAAACCGGTAG	7517
1858	TTTAACTACATACCTCATTTAGACCATTAACATAACATCAAAATTTGAGTGGTTTGACGAAGG	1917
7518	TTTAACTACATACCTCATTTAGACCATTAACATAACATCAAAATTTGAGTGGTTTGACGAAGG	7577
1918	CTTTTATGAGGCACCTAAAGGGGTATCTCTTGAGGATCTTTTGGCGACTGTCAAGTACTA	1977
7578	CTTTTATGAGGCACCTAAAGGGGTATCTCTTGAGGATCTTTTGGCGACTGTCAAGTACTA	7637
1978	TGTGGAACATCCAAACGAACGTCGGCATTCAGATTAATCGTTTGGTAAACGCTAGCGACCA	2037
7638	TGTGGAACATCCAAACGAACGTCGGCATTCAGATTAATCGTTTGGTAAACGCTAGCGACCA	7697
2038	TGTTCAAGAGAAACAAAATGGTCAAGCTGATACCAATCAAAACGGGAAAACCCAGCGAGGA	2097
7698	TGTTCTGTAATAAATAAGGTGAGACCAACAGTAAACCTCATGAAGATTAAGGAACATGATGA	7757
2098	GMAACCTTCAGACAGAAAACCTTGAGGAACA-----AACCCC2133	
7758	AGTAAGTGCACCACTCAACCTTGATCTGATGAAAAGAGAAATCACCGTGGTTTAAATCC	7817
2134	TCGAGAGAGAAAACCGCAAGCGGAGAAACAGAGTCTCCAAAACCAACAGAGGAACACAGA	2193
7818	TTACAGCAGATAATCTTTTATAAACCAAGCACTGATACGGAAAGAGACAGAGGAAGAGCTGA	7877
2194	AGAAATACACAGAGGAATCAGAGAAACCTCAGGTGCAGACTGAAAAGGTTGAGAGAAAACCT	2253
7878	AGATATCACAGATGAGGCTGAAATTCCTCAAGTAGAGAAATCTCTGTATTAAACGCTAAGAT	7937
2254	GAGHAGGGCTGAAGATTACTTTGGAAAAATCCAGAT2290	
7938	AGCAGATGCGGAGGCGCTTGTCTAGAAAAGTAAACAGAT7974	

SULT 11  
-09-884-465A-3

Sequence 3, Application US/09884465A

Publication No. US20030077293A1

**GENERAL INFORMATION:**

APPLICANT: Shire Biochem, Inc.

APPLICANT: Hamel, Jovsee

APPLICANT: Brodeur, Bernard.

APPLICANT: Martin, Denis

835 AGTCAAGCAAGTCAAGTAATGACATTGATAGTCTCTTGAACAGCTCTACAAACCTGCCT 894  
797 TTATCAGAACGCCATGTGGAATCTGATGGCCCTTATTTTCGACCGACGCAAAATCACAAGT 856  
895 TTGAGTCBAAGCCATGTAGAACTCTGATGGCCCTTATTTTCGACCGACGCAAAATCACAAGT 954  
857 CGAACCGCGAGAGGTAGTGTCTCCCTCATGGTAAACCAATACCACTTTATCCCTTATGAA 916  
955 CGAACCGCGAGAGGTAGTGTCTCCCTCATGGTAAACCAATACCACTTTATCCCTTATGAA 1014  
917 CAATGTCTGAAATTTGAAAAACGAATTCCTCGTATTTATCCCTTCGTATCGTTCAAAAC 976  
1015 CAATGTCTGAAATTTGAAAAACGAATTCCTCGTATTTATCCCTTCGTATCGTTCAAAAC 1074  
977 CATTTGGGTACCAAGTTCAGAACCAAGAACCACTCCACCAATCGACTCCGGAACCTAGT 1036  
1075 CATTTGGGTACCAAGTTCAGAACCAAGAACCACTCCACCAATCGACTCCGGAACCTAGT 1134  
1037 CCAAGTCGCAACCTGCAACCAAACTCTCAACGAGCTCCAAAGCAATCCCAATGATGAGAAA 1096  
1135 CCAAGTCGCAACCTGCAACCAAACTCTCAACGAGCTCCAAAGCAATCCCAATGATGAGAAA 1194  
1097 TTGGTCAAGAGAGCTGTTGAAAAAGTAGGCGATGTTATGTTCTTTTGGAGAGAAATCGAGTT 1156  
1195 TTGGTCAAGAGAGCTGTTGAAAAAGTAGGCGATGTTATGTTCTTTTGGAGAGAAATCGAGTT 1254  
1157 TCTCGTTATATCCAGCGCAAGAGTCTTTTTCAGCAGAAAACAGCAGCAGGCAATGATAGCAAA 1216  
1255 TCTCGTTATATCCAGCGCAAGAGTCTTTTTCAGCAGAAAACAGCAGCAGGCAATGATAGCAAA 1314  
1217 CTGGCCCAAGCAGGAAAGTTTATCTCAVAAAGCTAGGAGCTAAGAAAACTGACCTCCATCT 1276  
1315 CTGGCCCAAGCAGGAAAGTTTATCTCAVAAAGCTAGGAGCTAAGAAAACTGACCTCCATCT 1374  
1277 AGTGATCGAGAAATTTTCAATPAAGCTTATGACTTACTAGCAAGAAATTCACCAAGATT 1336  
1375 AGTGATCGAGAAATTTTCAATPAAGCTTATGACTTACTAGCAAGAAATTCACCAAGATT 1434  
1337 CTGGATTAATTAAGTTCGACAGATGATTTTGGAGCTTTGGATTAACCTGTTGAAAGCACTC 1396  
1435 CTGGATTAATTAAGTTCGACAGATGATTTTGGAGCTTTGGATTAACCTGTTGAAAGCACTC 1494  
1397 AAGGATGTCNCAAGTGATAAAGTCAAGTGTAGTGGAGATATTTCTTGCTCTCTTACTCGG 1456  
1495 AAGGATGTCNCAAGTGATAAAGTCAAGTGTAGTGGAGATATTTCTTGCTCTCTTACTCGG 1554  
1457 ATTGCTCATCCAGAACGTTTATAGGAAAAACCAATCGCAAAATTCACCTACACTGATGATGAG 1516  
1555 ATTGCTCATCCAGAACGTTTATAGGAAAAACCAATCGCAAAATTCACCTACACTGATGATGAG 1614  
1517 ATTCAAGTAGCCAAATTTGGCAGGCAAGTACACACAGAAAGCGTTTATCTTTGATCCT 1576  
1615 ATTCAAGTAGCCAAATTTGGCAGGCAAGTACACACAGAAAGCGTTTATCTTTGATCCT 1674  
1577 CGTGATATAACCACTGATGAGGCGGATGCCCTGATGTAACCTCCACATATGACCAATAGCCAC 1636  
1675 CGTGATATAACCACTGATGAGGCGGATGCCCTGATGTAACCTCCACATATGACCAATAGCCAC 1734  
1637 TGGATTAATAAAGATAGTTTCTGTAAGCTGAGAGCGGCGAGCCAGGCTTATGCTAAA 1696  
1735 TGGATTAATAAAGATAGTTTCTGTAAGCTGAGAGCGGCGAGCCAGGCTTATGCTAAA 1794  
1697 GAGAAAGGTTTACCCCTCTTCAGACAGCATCAGGATTCAGGAATTAATCTGAGGCAAAA 1756  
1795 GAGAAAGGTTTACCCCTCTTCAGACAGCATCAGGATTCAGGAATTAATCTGAGGCAAAA 1854  
1757 GGACGAGAGCTATCTCAACCGCGGTGAAGCAGCTAAGAGGTCCTTCTGATCGTATG 1816  
1855 GGACGAGAGCTATCTCAACCGCGGTGAAGCAGCTAAGAGGTCCTTCTGATCGTATG 1914  
1817 CTTTACCAATCTTCAATATCTGTAGAGAGTCAAAAACGGTAGTTTAAATCATACCTCATTAT 1876

1915 CTTTCAATCTTTCAATATATCTGTAGAAGTCAAAAACGGTAGTTTAAATCATACCTCATTAT 1974  
1877 GACCAATTACCAATAACATCAAAATTTGAGTGGTTTTCGAAAGGCTTTTATGAGGACCTAAG 1936  
1975 GACCAATTACCAATAACATCAAAATTTGAGTGGTTTTCGAAAGGCTTTTATGAGGACCTAAG 2034  
1937 GGGTATATCTTTGAGGATCTTTTGGCGACTGTCTCAAGTACTATGTCGAACATCCAAACGAA 1996  
2035 GGGTATATCTTTGAGGATCTTTTGGCGACTGTCTCAAGTACTATGTCGAACATCCAAACGAA 2094  
1997 CGTCCGCAATCAGATTAATGTTTGGTAAAGCTAGGACCATGTTTCAAGAAACAAAAAT 2056  
2095 CGTCCGCAATCAGATTAATGTTTGGTAAAGCTAGGACCATGTTTCAAGAAACAAAAAT 2154  
2057 GGTCAAGCTGTATACCAATCAAAACGGAACCAAGCGAGGAGAAACCTCAGACAGAAAA 2116  
2155 GGTCAAGCTGTATACCAATCAAAACGGAACCAAGCGAGGAGAAACCTCAGACAGAAAA 2214  
2117 CTTGAGGAGAGAAACCCCTCGAGAGAGAGAAACCCGCAAGCGAGAGAAACCTCAGAGTCTCCAAA 2176  
2215 CTTGAGGAGAGAAACCCCTCGAGAGAGAGAAACCCGCAAGCGAGAGAAACCTCAGAGTCTCCAAA 2274  
2177 CCAACAGAGGAAAC---AGAAGATCAACAGAGAAATCAGAGAAACCTCAGGTCGAGACT 2233  
2275 CCAACAGAGGAAACCCGAGAGAGAAATCACCAGAGGAAATCAGAGAAACCTCAGGTCGAGACT 2334  
2234 GAAAAGGTTGAAGAAAAACCTGAGAGAGGCTGAAGATTTTACTTGGAAAAATCCAGGAT 2290  
2335 GAAAAGGTTGAAGAAAAACCTGAGAGAGGCTGAAGATTTTACTTGGAAAAATCCAGGAT 2391

## RESULT 12

US-09-884-465A-4  
; Sequence 4, Application US/09884465A  
; Publication No. US2003007293A1  
; GENERAL INFORMATION:  
; APPLICANT: Shire Biochem, Inc.  
; APPLICANT: Brodeur, Bernard  
; APPLICANT: Martin, Denis  
; APPLICANT: Charland, Nathalie  
; APPLICANT: Ouellet, Catherine  
; TITLE OF INVENTION: Streptococcus Antigens  
; FILE REFERENCE: 055190-0044  
; CURRENT APPLICATION NUMBER: US/09/884,465A  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: 60/212,683  
; NUMBER OF SEQ ID NOS: 384  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 2647  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-09-884-465A-4

Query Match 81.1%; Score 1856.8; DB 10; Length 2647;

Best Local Similarity 89.0%; Pred. No. 0;

Matches 2079; Conservative 0; Mismatches 207; Indels 51; Gaps 5;

QY 2 TGTCTCTATGAACCTTGGTCGTCCACCAAGCTGGTTCAGGTTTAAAGAAAGTCTAATCCGAGTT 61  
DB 102 TGTGCTTATGAACCTTGGTCGTCCACCAAGCTGGTTCAGGTTTAAAGAAAGTCTAATCCGAGTT 158  
QY 62 TCTTATATAGATGGTGTGATCAGGCTGGTCAAAAAGCGCAAAAACTTGACACCAAGTGAAGTC 121  
DB 159 TCTTATATAGATGGTGTGATCAGGCTGGTCAAAAAGCGCAAAAACTTGACCTCTGATGAGTT 218  
QY 122 AGTAAGAGGAGGAGGATCAACCGCGAAACAAATNGTNATCAAGATTTCGATTCAGGTTAT 181  
DB 219 AGCAAGCGTGAAGGAAATCAACCGCGAAACAAATTCGTCATCAAGATTACGGATCAAGTTAT 278  
QY 182 GTGACCTCTCTGAGAGACCAATTTATCTACTATATGCGCAAGGTTCTTATGATGCCATC 241

279 GTGACCTCTCATGGAGACCATTTATCATTTACTAATATGGCAAGTCCCTTATGATGCCATC 338  
242 ATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTATCAGTTGAGGATTCAGACATTGTC 301  
339 ATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTATCAGTTGAGGATTCAGACATTGTC 398  
302 AATGAAATCAAGGGTGGTTATGTCTATTAAGGTAAACGGTAAATATCTATGNTACCTTAAG 361  
399 AATGAAATCAAGGGTGGTTATGTCTATTAAGGTAAACGGTAAATATCTATGNTACCTTAAG 458  
362 GATGACGCTCATGCGGATTAATTTGGAACAAAGAGAGATTAACCGTCAGAGCAGGAA 421  
459 GATGACGCTCATGCGGATTAATTTGGAACAAAGAGAGATTAACCGTCAGAGCAGGAA 518  
422 CGCAGTCATATCAT-----AACTCAAGAGCAGATAATGCTGTTGCTGCGCCAGA 472  
519 CATAGTCAGCATCGTGAAGAGGGACTTCAGCAAAACGATGGTGGGTAGCCTTTGCAAGT 578  
473 GCCCAAGGACGTTATACCAAGGATGATGGGTATATCTTCAATGCACTCTGATATCATTTGAG 532  
579 TCACAGGGAAGCTACACACAGATGATGGTTATATCTTCAATGCACTCTGATATCATTCGAA 638  
533 GACACGGGTGATCTTATATCCTTCTCACGGGACCAATTAACCATTAATTCCTTAAGAAT 592  
639 GATACGGGCGATCTATATCCTTCTCATGGAGATCATTAACCATTAATTCCTTAAGAAT 698  
593 GAGTTATCAGCTAGCAGTACTGCTGACAGAGCCTATGGAATGGGAAGCAGGATCT 652  
699 GAGTTATCAGCTAGCAGTACTGCTGACAGAGCCTTCTATCTGCTGGGAAAAATCTG 758  
653 CGTCCCTT-----CTTCAAGTTCTAGTTATAAT 679  
759 TCAAATTTAAGAACCTATGCGCGACAAATAGGATPAACTCCCAAGAACAACTGGGTA 818  
680 GCAATCAGCTCAACCAAGATTGTACAGAAACCAATCTGACTGTCACTCCAACTTAT 739  
819 CTTCTGTAAGCAATCCAGGAATACTAATACTAACACAAAGCAACACCACTAAC 878  
740 CATCA---AATCAAGGGGAAACATTTCAAGCCTTTTACGTGAATGTATGTAACCC 796  
879 AGTCAAGCAAGTCAAAAGTAATGACATTTGATAGTCTCTTGAAACAGCTCTACAACTGCT 938  
797 TTATCAGAAAGCCATGTGAATCTGATGCGCTTATTTTCGACCCAGCGCAATCAAGT 856  
939 TTGAGTCAAGCCATGTGAATCTGATGCGCTTATTTTCGACCCAGCGCAATCAAGT 998  
857 CGAACCGCCAGAGGTAGTCTGCTCATGTATACCAATTAACCACTTTATCCCTTATGAA 916  
999 CGAACCGCCAGAGGTAGTCTGCTCATGTATACCAATTAACCACTTTATCCCTTATGAA 1058  
917 CAATGTCTGAATTGGAAAAAGAAATGCTGCTATTTATTTCCCTTATCGTTCAAAAC 976  
1059 CAATGTCTGAATTGGAAAAAGAAATGCTGCTATTTATTTCCCTTATCGTTTCAAAAC 1118  
977 CATTCGGTACAGATTCAGACCAAGAACCAACCAAGTCCCAATTCGATCCGGAACCTAGT 1036  
1119 CATTCGGTACAGATTCAGACCAAGAACCAACCAAGTCCCAATTCGATCCGGAACCTAGT 1178  
1037 CCAAGTCCGCAACCTGCACCAATCTCAACAGCTCCCAAGCAATCCAAATGATGAGAAA 1096  
1179 CCAAGTCCGCAACCTGCACCAATCTCAACAGCTCCCAAGCAATCCAAATGATGAGAAA 1238  
1097 TTGGTCAAGAGAGCTGTTTCGAAAAGTAGCGATGTTATGCTTTGAGGAGATGGAGTT 1156  
1239 TTGGTCAAGAGAGCTGTTTCGAAAAGTAGCGATGTTATGCTTTGAGGAGATGGAGTT 1298  
1157 TCTCGTTATATCCGACCAAGAGATCTTTTCAGCAGAAAACAGCAGCAGGCAATGATGACAAA 1216  
1299 TCTCGTTATATCCGACCAAGAGATCTTTTCAGCAGAAAACAGCAGCAGGCAATGATGACAAA 1358  
1217 CTGGCCACGAGGAAGTTTATCTCATAGCTAGGAGCTAAGAAAACAGCTCCCATCT 1276

1359 CTGCCAAGCAGGAAAGTTTATCTCATAAGCTAGGAGCTAAGAAAACAGCTCCCATCT 1418  
1277 AGTGATCGAGAAATTTTACAATAAGGCTTATCACTTACTAGCAAGAAATTCACCAAGATTTA 1336  
1419 AGTGATCGAGAAATTTTACAATAAGGCTTATCACTTACTAGCAAGAAATTCACCAAGATTTA 1478  
1337 CTTGATAATAAAGTTCGACAAAGTTGATTTTGAGGCTTTGGATTAACCTGTTGGAACGACTC 1396  
1479 CTTGATAATAAAGTTCGACAAAGTTGATTTTGAGGCTTTGGATTAACCTGTTGGAACGACTC 1538  
1397 AAGGATCTCNCAGTATAAAGTCAAGTTAGTGGANGATATTTCTTGCCCTTTTACTGCTCG 1456  
1539 AAGGATCTCNCAGTATAAAGTCAAGTTAGTGGANGATATTTCTTGCCCTTTTACTGCTCG 1598  
1457 ATTTCGTATCCAGNAACGTTTAGGAAAACCAATGGGAAAATTTACCTACACTGATGAG 1516  
1599 ATTTCGTATCCAGNAACGTTTAGGAAAACCAATGGGAAAATTTACCTACACTGATGAG 1658  
1517 ATTCAAGTAGCAGGTTGGCAGGCAAGTACACAACAGAAAGCGGTTATCTTTGATCCT 1576  
1659 ATTCAAGTAGCAGGTTGGCAGGCAAGTACACAACAGAAAGCGGTTATCTTTGATCCT 1718  
1577 CGTGATAATAACCAAGTATGAGGGGATGCTTATGTAACTCCACATATGACCCATAGCCAC 1636  
1719 CGTGATAATAACCAAGTATGAGGGGATGCTTATGTAACTCCACATATGACCCATAGCCAC 1778  
1637 TGAATTTAAAAAGATAGTTTGTCTGAAAGCTGAGAGAGCGGAGCCAGGCTTTATGCTAAA 1696  
1779 TGAATTTAAAAAGATAGTTTGTCTGAAAGCTGAGAGAGCGGAGCCAGGCTTTATGCTAAA 1838  
1697 GAGAAAGGTTTGACCCCTCCTTCGACAGACCATCAGGATTCAGGAAATACCTGAGGCAAAA 1756  
1839 GAGAAAGGTTTGACCCCTCCTTCGACAGACCATCAGGATTCAGGAAATACCTGAGGCAAAA 1898  
1757 GGAGCAGAGCTATCTACAAACCGGCTGAAAGCAGCTAAGAAAGTGCCTTCTGATCGTATG 1816  
1899 GGAGCAGAGCTATCTACAAACCGGCTGAAAGCAGCTAAGAAAGTGCCTTCTGATCGTATG 1958  
1817 CCTTACATCTTCAATATACCTGTAGAGAGTCAAAAACGCTAGTTTAACTCATACCTCATAT 1876  
1959 CCTTACATCTTCAATATACCTGTAGAGAGTCAAAAACGCTAGTTTAACTCATACCTCATAT 2018  
1877 GACCATTTACCATTAACATCAAAATTTGAGTGGTTTTCGAAAGGCTTTTATGAGGCACTAAG 1936  
2019 GACCATTTACCATTAACATCAAAATTTGAGTGGTTTTCGAAAGGCTTTTATGAGGCACTAAG 2078  
1937 GGTATATCTCTTGAGGATCTTTTGGCGACTGTCAAGTACTATGCGAACATCCAAACGAA 1996  
2079 GGTATATCTCTTGAGGATCTTTTGGCGACTGTCAAGTACTATGCGAACATCCAAACGAA 2138  
1997 GGTCCGCAATTCAGATAATGTTTGGTAAACGCTAGCGACCATGTTTCAAGAAACAAAAAT 2056  
2139 GGTCCGCAATTCAGATAATGTTTGGTAAACGCTAGCGACCATGTTTCAAGAAACAAAAAT 2198  
2057 GGTCAAGCTGATACCAATCAAAACGGAACCAACGAGGAGGAAACCTCAGACAGAAAA 2116  
2199 GGTCAAGCTGATACCAATCAAAACGGAACCAACGAGGAGGAAACCTCAGACAGAAAA 2258  
2117 CTTGAGGAGAAACCCCTTCGAGAGAGAAACCGCAAAACGAGGAGAAACAGAGTCTCCAAAA 2176  
2259 CTTGAGGAGAAACCCCTTCGAGAGAGAAACCGCAAAACGAGGAGAAACAGAGTCTCCAAAA 2318  
2177 CCAACAGAGAAACCC---AGAAGATCAACAGAGGAAATCAGAAAGAACCTCAGGTCGAGACT 2233  
2319 CCAACAGAGAAACCCAGAGAGAAATCACCAGAGGAAATCAGAAAGAACCTCAGGTCGAGACT 2378  
2234 GAAAGGTTGAGAAACCTGAGAGAGGCTGAAATTTACTTTGGAATAATCCAGAT 2290  
2379 GAAAGGTTGAGAAACCTGAGAGAGGCTGAAATTTACTTTGGAATAATCCAGAT 2435

RESULT 13

US-10-158-844-243

; Sequence 243, Application US/10158844  
; Publication No. US20040029118A1  
; GENERAL INFORMATION:  
; APPLICANT: Kunsch et al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-R  
; COMPUTER: Dell Latitude Pentium 3  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/158,844  
; FILING DATE: 03-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/961,527  
; FILING DATE: 1997-10-30  
; APPLICATION NUMBER: US 60/029,960  
; FILING DATE: 1996-10-31  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hyman, Mark J.  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB340PDI  
; INFORMATION FOR SEQ ID NO: 243:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2359 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 243:  
US-10-158-844-243  
  
Query Match 60.1%; Score 1376.6; DB 13; Length 2359;  
Best Local Similarity 97.8%; Pred. No. 0;  
Matches 1410; Conservative 0; Mismatches 14; Indels 18; Gaps 1;  
  
QY 1 TTGTTCTATGAACCTGGTGGTCAACCAAGCTGGTCAAGGTTAAGAAAGAGTCTAATCCAGT 60  
Db 936 TTGTTCTATGAACCTGGTGGTCAACCAAGCTGGTCAAGGTTAAGAAAGAGTCTAATCCAGT 995  
  
QY 61 TTCTTATATAGATGGTATCAGGCTGGTCAAAAGCGCAAAACTTTGACACAGATGAAGT 120  
Db 996 TGCTTATATAGATGGTATCAGGCTGGTCAAAAGCGCAAAACTTTGACACAGATGAAGT 1055  
  
QY 121 CAGTAAGAGGAGGGGATCAACCGCAACCAATNGTATCAAGATTACGGATCAAGTTA 180  
Db 1056 CAGTAAGAGGAGGGGATCAACCGCAACCAATNGTATCAAGATTACGGATCAAGTTA 1115  
  
QY 181 TGTGACCTCTCATGGAGACCAATTATCAATTAATGGAAGGTTCTCTATGATGCCAT 240  
Db 1116 TGTGACCTCTCATGGAGACCAATTATCAATTAATGGAAGGTTCTCTATGATGCCAT 1175  
  
QY 241 CATCAGTGAAGAGCTCTCTATGAAAGATCCGAATTTATCAGTTGAAGGATTCAGACATGTT 300  
Db 1176 CATCAGTGAAGAGCTCTCTATGAAAGATCCGAATTTATCAGTTGAAGGATTCAGACATGTT 1235  
  
QY 301 CAATGAATCAAGGTTGGTTATGTCATTAGGTTAAACGGTAAATACATGATGTTACCTTAA 360  
Db 1236 CAATGAATCAAGGTTGGTTATGTCATTAGGTTAAACGGTAAATACATGATGTTACCTTAA 1295  
  
QY 361 GGATCAGCTCATGCGGATTAATTCGACAAAAGAGAGATTAAAGCTCAGAGCAGGA 420  
Db 1296 GGATCAGCTCATGCGGATTAATTCGACAAAAGAGAGATTAAAGCTCAGAGCAGGA 1355  
  
QY 421 ACGCAGTCATAATCAATACTCAAGAGCAGATAATGCTGTTGCTGCAGCCAGAGCCCAAGG 480

RESULT 14

Db 1356 ACGCAGTCATAATCAATACTCAAGAGCAGATAATGCTGTTGCTGCAGCCAGAGCCCAAGG 1415  
QY 481 ACGTTATACAACGGATGATGGGTATATCTTCAATGATCTGATATCATTTGAGGACACGGG 540  
Db 1416 ACGTTATACAACGGATGATGGGTATATCTTCAATGATCTGATATCATTTGAGGACACGGG 1475  
QY 541 TGATGCTTATATCGTTCTCTCAGCGGACCAATTAACCATTAATCCATTAATCCATTAAGATGATGTTATC 600  
Db 1476 TGATGCTTATATCGTTCTCTCAGCGGACCAATTAACCATTAATCCATTAAGATGATGTTATC 1535  
QY 601 AGCTAGCGAGTTAGCTGCTGCAGAGACCCCTATTGGAATGGGAAGCAGGATCTGCTCTTC 660  
Db 1536 AGCTAGCGAGTTAGCTGCTGCAGAGACCCCTATTGGAATGGGAAGCAGGATCTGCTCTTC 1595  
QY 661 TTCAAGTCTTAGTTATAATGCAATCCAGCTCAACCAAGATTGTCAGAGAACCAACAATCT 720  
Db 1596 TTCAAGTCTTAGTTATAATGCAATCCAGCTCAACCAAGATTGTCAGAGAACCAACAATCT 1655  
QY 721 GACTGTCACCTCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGA 780  
Db 1656 GACTGTCACCTCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGA 1715  
QY 781 ATTGTATGCTAAACCTTATCAGAACCGCATGFGGAATCTGATGGCTTTATTTTGAACC 840  
Db 1716 ATTGTATGCTAAACCTTATCAGAACCGCATGFGGAATCTGATGGCTTTATTTTGAACC 1775  
QY 841 AGCGCAAAATCACAAGTCGAAACCGCCAGAGGTGTAGCTGCTCCCTCATGTTAACCATTACCA 900  
Db 1776 AGCGCAAAATCACAAGTCGAAACCGCCAGAGGTGTAGCTGCTCCCTCATGTTAACCATTACCA 1835  
QY 901 CTTTATCCCTTATGAACAAATGTCGAAATGGAAAAACGAATTCGTGATTTATTTCCCT 960  
Db 1836 CTTTATCCCTTATGAACAAATGTCGAAATGGAAAAACGAATTCGTGATTTATTTCCCT 1895  
QY 961 TCGTTATCGTTCAAAACCAATGCGGTACCAAGTCAAGACCAAGAACCAAGTCCACAATC 1020  
Db 1896 TCGTTATCGTTCAAAACCAATGCGGTACCAAGTCAAGACCAAGAACCAAGTCCACAATC 1955  
QY 1021 GACTCCGGAACCTAGTCCAAAGTCGCAACCTGCACCAAAATCTCAACCAAGTCCCAAGCAA 1080  
Db 1956 GACTCCGGAACCTAGTCCAAAGTCGCAACCTGCACCAAAATCTCAACCAAGTCCCAAGCAA 1997  
QY 1081 TCCAAATTGATGAGAAATTTGGTCAAAAGAGCTGTTGAAAAAGTAGGCGATGTTATGTTT 1140  
Db 1998 TCCAAATTGATGAGAAATTTGGTCAAAAGAGCTGTTGAAAAAGTAGGCGATGTTATGTTT 2057  
QY 1141 TGAGGAGAAATGAGTTTCTCGTTTATATCCAGCAAGGATCTTTTCAGAGAAAACAGCAGC 1200  
Db 2058 TGAGGAGAAATGAGTTTCTCGTTTATATCCAGCAAGGATCTTTTCAGAGAAAACAGCAGC 2117  
QY 1201 AGGCATTGATCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAA 1260  
Db 2118 AGGCATTGATCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAA 2177  
QY 1261 AACTGACCTCCCATCTAGTGTATGAGAAATTTTACAAATAGGCTTATGACTTACTAGCAAG 1320  
Db 2178 AACTGACCTCCCATCTAGTGTATGAGAAATTTTACAAATAGGCTTATGACTTACTAGCAAG 2237  
QY 1321 AATTCAACAAGATTTACTGTAATAAAGGTCGAACAAGTTGATTTTGAAGCTTTGGATAA 1380  
Db 2238 AATTCAACAAGATTTACTTTGATAATAAAGGTCGAACAAGTTGATTTTGAAGCTTTGGATAA 2297  
QY 1381 CCTGTTGGAACGACTCAAGGATGTCNCAAGTGAATAAAGTCAAGTTAGTGGANGATTTCT 1440  
Db 2298 CCTGTTGGAACGACTCAAGGATGTCNCAAGTGAATAAAGTCAAGTTAGTGGANGATTTCT 2357  
QY 1441 TG 1442  
Db 2358 TG 2359



US-10-412-862-9  
; Sequence 9, Application US/10412862  
; Publication No. US20040052781A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; TITLE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-685  
; CURRENT APPLICATION NUMBER: US/10/412,862  
; CURRENT FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 09/468,656  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 9  
; LENGTH: 2451  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(2451)  
; OTHER INFORMATION: n = a, c, t or g  
US-10-412-862-9

Query Match 43.3%; Score 990.6; DB 13; Length 2451;  
Best Local Similarity 67.8%; Pred. No. 1.3e-267;  
Matches 1484; Conservative 0; Mismatches 645; Indels 60; Gaps 5;  
1 TRGTTCCTATGAACTGGTGGCTGCTCACCAGCTGCTCAGGTTTAAAGAGTCTTAATCAGT 60  
57 TGTGTTCTTACGAGTGGGACTGTATCAGCTGAACGGTTTAAAGAAA---TAACTCGTGT 113  
61 TCTTATATAGATGGTGTGATCAGGCTGGTCAAAAGGCGAGAAAACCTTGACACAGATGAAGT 120  
114 TTCCTATATAGATGGAACAAACAGCGACGCAAAACCGAGAAATTTGACTCTCTGATGAGT 173  
121 CAGTAAGAGGGAGGGATCAACGCGCGAAACAAATGNGTATCAAGATTCAGGATCAAGGTTA 180  
174 TAGCAAGCGTGAAGGAATCAATGCTGAGCAATCGTCATCAAGATAACAGCAAGGCTA 233  
181 TGTGACCTCTCATGAGACCATATATCAATTAATGCAAGGTTCCCTTATGATGCCAT 240  
234 TGTCACTTCACATGGCGACCATATCATTTATTAACAATGGTAAGGTTCCCTTATGACGCTAT 293  
241 CATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTTATCAGTTCAAGGATTCAGACATTTG 300  
294 CATCAGTGAAGAAATTAATCATGAAAGATCCAACTATAGCTTAAAGATGAGATATTGT 353  
301 CAATGAATCAAGGTGGTATGTCATTAAGGTAACGGTAAACGTTAAATATCTATGNTACCTTAA 360  
354 TAATCAGGTCAAGGGTGGATATGTTATCAAGGTAGATGGAAAATTAATGTTTACCTTAA 413  
361 GGATCAGCTCATGCGGATAATATTCGGCAAGAGAGAGATTAAGAGTTCAGAGCAGGA 420  
414 GGATCCTGCCACGGGATAACGTCGCTGTAAGAGGAAATCAATCGACAAACAAACGA 473  
421 ACAGAGTCATATATCAT-----AACTCAAGAGCAGATATGCTGTTGCTGCAGCCAG 471  
474 GCAATGTCACATCGTAGGTGGAACTCCAGAAACGATGGTGGCTGTGCTTGGCAAG 533  
472 AGCCCAAGACGTTATACAAACGGATGATGGGTATATCTTCAATGCACTGATATCATTTGA 531  
534 TTGCAAGAGCAGCTATACAGATGATGTTTATATCTTTTAAATGCTTCTGATATCATAGA 593  
532 GGACAGCGGTGATGCTTATATGTTTCCCTCAGCGGACCATTTACCAATTAATTCCTAAGAA 591  
594 GGATCTGTTGATGCTTATATGTTTCCCTCATGAGATCATTTACCAATTTACCTAAGAA 653  
592 TGAGTTATCAGCTAGCGATTAGCTGCTGCAGAGCCT----- 629

Db 654 TGAGTTATCAGCTAGCGAGTTGGCTGCTGCAGAGCCTTCCCTATCTGGTCGAGGAAATCT 713  
QY 630 -----ATTGGAATGGAAGCAGGAGTCTGGTCTCTTCAAGTTCTAGTTATATAA 678  
Db 714 GTCAAAATCAAGAACCTTATCGCCGACAAATAGCGATAACACATCTCAAGAACAAACTGGGT 773  
QY 679 TCGAAATCCAGCTCAACCAAGATTGTCAGAGAACCAATCTGACTGTCACCTCAACTTA 738  
Db 774 ACCTTCTGTAAGCAATCCAGAACCTAACAATATCTAACCAAGCAACAAGCAACTAA 833  
QY 739 TCATCA---AAATCAAGGGGAAAAACATTTCAAGCCCTTTTACGTGAATTTGATCTTAAACC 795  
Db 834 CAGTCAAGCAAGTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 893  
QY 796 CTTATCAGAACCCCATGATGGAATCTGATGGCCCTTATTTTCGACCCAGCGCAATCAAG 855  
Db 894 TTTGAGTCAACGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 953  
QY 856 TCGAACCGCAGAGGTGATGCTCCCTCATGTAACCATTTACCATTTTATCCCTTATGA 915  
Db 954 TCGAACAGCTAGAGGTGTCAGTGGCCACACGAGATCAATTACCATTTATCCCTTACTC 1013  
QY 916 ACAAATGCTCAATTTGGAATAACGAATTTGCTGATTTATTTCCCTTCTGTTTATCGTTCAA 975  
Db 1014 TCAATGCTCGAATTTGGAAGAACGATGCTGATTTATTTCCCTTCTGTTTATCGTTCAA 1073  
QY 976 CCATTGGGTACAGATTCAAGACCAAGAACCAAGTCCCAATCGACTCCGGAACCTAG 1035  
Db 1074 CCAATGGGTACAGATTCAAGGCCAGAACCAAGTCCCAAGTCCCAAGTCCGGAACCTAG 1133  
QY 1036 TCCAGTCCGCAACCTGACCAATCTTCAACAGTCCCAAGTCCCAAGTCCCAAGTCCCAAG 1095  
Db 1134 TCCAGGCCCGCAACCTGACCAATCTTAAATAGACTCAA-----ATTCTTC 1181  
QY 1096 ATTGCTCAAGAAAGCTGTTTCAAGAAAGTAGGCGATGTTATGCTTTTGAGAGAAATCGAGT 1155  
Db 1182 TTTGTTAGTCACTGTTGACGAAAGTTGGGAGAGATGATTTTGAAGAAAGAGGCAT 1241  
QY 1156 TTTCTGTTATATCCAGCAAGGATCTTTTCAAGCAAGAACAGCAGCAGGCAATTTGATAGCAA 1215  
Db 1242 CTCTCGTTATGCTTTGCAAGAAATTTACCATCTGAAACTGTTAAATACTTTGAAGCAA 1301  
QY 1216 ACTGGCCAGCAGAGAAATTTATCTATAAGCTAGAGCTTAAGAAAACGTGACCTCCCATC 1275  
Db 1302 GTTATCAAAACAGAGAGTGTTCACACACTTTTAACTGCTTAAAGAAAGAAATTTGTGCTCC 1361  
QY 1276 TAGTGATCGAGAAATTTTCAATAAAGGCTTTATGACTTCTAGCAAGAAATTCACCAAGATTT 1335  
Db 1362 TCGTCAAGCAAGAAATTTTATGATAAGCAATATCTGTTAACTGAGGCTCATAAAGCCTT 1421  
QY 1336 ACTTGATTAATAAGGTCGACAAAGTTGATTTTGGAGCTTTGGATTAACCTGTTGGAACGCT 1395  
Db 1422 GTTTGNAATAAAGGTCGTAACTCTGATTTTCAAGCCTTAGACAAATTTATAGAAGCCTT 1481  
QY 1396 CAAGGATGTCNCAAGTGAATAAGTCAAGTTAGTGAAGATATTTTGGCTTCTTATGCTCC 1455  
Db 1482 GAATGATGAATCGACTAATAAGAAAAATTTGTTAGATGATTTTGGCATTTCTTAGCACC 1541  
QY 1456 GATTTCGTCATCCAGAACGTTTAGAAAACCAAAATCGCAAAATTTACCTACATGATGATGA 1515  
Db 1542 AATTACCATCCAGAGCGACTTGGCAACCAAAATTTCAAATTTGAGTATATGAAAGAGA 1601  
QY 1516 GATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACAAAGAACGCTTATATCTTTTATGATCC 1575  
Db 1602 AGTTCTGTTTGTCTCAATTTAGCTGATAGTATACAACTGATGATGATGATGATGATGATGAT 1661  
QY 1576 TCGTGAATTAACAGTGAATGAGGGGATGCTATGTAATCTCCATATGATGATGATGATGATGATGAT 1635  
Db 1662 ACATGATATTAATCAGTGAAGGAGATGATATGTAACGCTCATATGGGCAATGATCA 1721  
QY 1636 CTGGATTAAAGAGATGTTTGTGTAAGCTGAGAGAGCGCAGCCAGGCTTATGCTTAA 1695

Db 1722 CTGGATGGAAAGATAGAGCTTTCTGATAGGAAAGATTGCGAGCTCAAGCCTAATCTAA 1781  
QY 1696 AGAGAAAGGTTTGACCCCTCTTCGACAGACCATCAGGATTCAAGAAATCTGAGGCAAA 1755  
Db 1782 AGAAAAAGGTATCCCTACCTCCATCTCCAGACGAGATGTTAAAGCAATCCAACTGGAGA 1841  
QY 1756 AGAGCAGAAGCTATCTTAACCCGCTGAAAGCAGCTTAAGAGGTGCCACTTGATCTGAT 1815  
Db 1842 TAGTGCAGCAGCTATTTTACAAATCGTGTGAAAGGGGAAAAACGAATTCCTCGTGGACT 1901  
QY 1816 GCTTACAACTCTCAATATCTGTAAGTCAAAAGCGGTAGTTTAACTCATACCTCAATTA 1875  
Db 1902 TCCATATATGTTGAGCATACAGTGAAGTTAAAGCGTAATTTGATTAATCTCTATAA 1961  
QY 1876 TGACCAATTACATAACATCAAAATTTGAGTGGTTTGACGAAGGCCCTTTATGAGGCACTAA 1935  
Db 1962 GGATCAATTACCAATATTAATTTGCTTGGTTTGATGATCACACATACAAAGCTCCAAA 2021  
QY 1936 GGGGTATCTCTTGAGGATCTTTTGGGACCTGTCAAGTACTATGTGAAACATCCAAACGA 1995  
Db 2022 TGGCTATACCTTGGAAAGATTTGTTTGGGACGATTTAAGTACTACGTAGAACAACCTGACGA 2081  
QY 1996 AGCTCCGCATTCAGATTAATGTTTGGTAAAGCTAGCGACCATGTTCAAAGAAAACAAAA 2055  
Db 2082 ACCTCCACATCTTAAATGATGATGGGCAATGCCAGTGAGCATGTGTTAGCAAGAAAGA 2141  
QY 2056 TGGTCAAGCTGTATACCAATCAAAACGGHAAACCAAGCGAGGAGAAACCTCAGACAGAAA 2115  
Db 2142 CCACAGTGAAGATCCAAATAAGAACTTCAAAGCGGATGAAGAGCCAGTAGAGAAACACC 2201  
QY 2116 ACTGAGGAAGAAACCCCTCGAGAAAGA 2144  
Db 2202 TGCTGAGCCAGAAGTCCCTCAAGTAGAGA 2230

## RESULT 15

US-10-412-850-9  
; Sequence 9, Application US/10412850  
; Publication No. US2004001836A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; TITLE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-686  
; CURRENT APPLICATION NUMBER: US/10/412,850  
; CURRENT FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 09/468,656  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 9  
; LENGTH: 2451  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(2451)  
; OTHER INFORMATION: n = a, c, t or g  
US-10-412-850-9

Query Match 43.3%; Score 990.6; DB 16; Length 2451;  
Best Local Similarity 67.8%; Pred. No. 1.3e-267;  
Matches 1484; Conservative 0; Mismatches 645; Indels 60; Gaps 5;  
QY 1 TTGTTCTCTATGAATCTGTCCTCCACGAGCTGGTCAAGGTTAAGAAAGAGTCTAATTCGAGT 60  
Db 57 TTGTTCTTACGAGTTGGGACTGTATCAAGCTAGAACGGTTAAGGAAA---TAATTCGTGT 113  
QY 61 TTCCTATATAGATGGTATCAGGCTGGTCAAAAGGCAAGAAACTTGACACCAGATGAAGT 120

Db 114 TTCCTATATAGATGAAGAACAGCGACGCAAAAACCGAGAATTTGACTTCCTGATGAGGT 173  
QY 121 CAGTAAGAGGAGGAGGATCAACGCCGAAACAAATNGTNATCAAGATTAACGATCAAGGTTA 180  
Db 174 TAGCAAGCGTGAAGGAATCAATGCTGAGCAATCGTCAAGATAACAGACCAAGGCTA 233  
QY 181 TGTGACCTCTCATGGAGACCAATATCATTAATATGSCAAGGTTCCCTTATGATGCCAT 240  
Db 234 TGTCACTTCAATGCGGACCACTATCATTTATCAATGATGAAGTTCCCTTATGACGCTAT 293  
QY 241 CATCAGTCAAGAGCTCCCTCATGAAGATCCGAATTTATCAGTTGAAGGATTCAGACATTTGT 300  
Db 294 CATCAGTGAAGATTTACTCATGAAGATCCAACTATATAGCTAAAGATGAGGATTTGT 353  
QY 301 CAATGAATCAAGGTTGTTATGTCATTAAAGTAAACGGTAAATATCTATGTTNATACCTTAA 360  
Db 354 TAATGAGGTCAAGGTTGATATGTTATCAAGGTAGATGGAATACTATGTTTACCTTAA 413  
QY 361 GGATGCAGCTCATGCGGATATATTCGACAAAGAGAGATTAACGTCAGAGCAGGA 420  
Db 414 GGATGCTCCCAACGCGGATAACGTCGTAACAAAGAGGAAATCAATCGACAAAAACAAGA 473  
QY 421 ACGCAGTCAATATCAT-----AACTCAAGAGCAGATAATGCTGTGCTGCAGCCAG 471  
Db 474 GCATAGTCAACATCGTGAAGGTGGAACCTCAAGAAACGATGGTGTGCTGTGCTTGGCAGC 533  
QY 472 AGCCCAAGGAGCTTATACAACGGATGATGGGTATATCTTCAATGCAATCTGATATCATTTGA 531  
Db 534 TTGCAAGGAGCGCTATACTACAGATGATGGTTATATCTTTAATGCTTCTGATATCATAGA 593  
QY 532 GGACACGGGTGATGCTTATATGCTTCCTCAGCGGACCAATACCATTACATTCCTTAAGAA 591  
Db 594 GGATACTGTTGATGCTTATATGCTTCTCATGGAGATCATTTACCATTAATTCCTTAAGAA 653  
QY 592 TGAGTTATCAGCTAGGAGTTAGCTGCTGCAGAGGCT----- 629  
Db 654 TGAGTTATCAGCTAGGAGTTGGCTGCTGCAGAGGCTTCTCTATCTGGTCGAGGAAATCT 713  
QY 630 -----ATTGGAATGGGAAGCAGGAGTCTCGTCTCTTCAAGTTTCTAGTTTATAA 678  
Db 714 GTCAAATTCAGAAACCTATCGCCGACAAATAGCGATAACACTTCAAGAAACAACTGGGT 773  
QY 679 TGCAAATCAGCTCAACCAAGATTTGTGAGAACCAATCTGACTGTCTACTCCAATTGA 738  
Db 774 ACCTTCTGTAGCAATCCAGGAACTACAATATCTACACAGCAACACAGCAACACTAA 833  
QY 739 TCATCA---AATCAAGGGGAAACATTTCAAGCCTTTTACGTGAATTTGTATGCTAAAC 795  
Db 834 CAGTCAAGCAAGTCAAGATTAATGACATGATGCTCTTTGAAACAGCTCTACAACTGCC 893  
QY 796 CTTATCAGAACGCCATGTGGAATCTGATGGCCTTATTTTCGACCCAGCGCAATTCACAG 855  
Db 894 TTTGAGTCAACGACATGATGATCTGATGGCCTTGTCTTTGATCCAGACAAATCAACAG 953  
QY 856 TCGAACCGCCAGAGGTGATGCTCCCTCATGTTAACCAATACCATTATTCCTTATGA 915  
Db 954 TCGAACAGCTAGAGGTGTTGACGTGCCACACGAGAGATCAATTACCACTTCATCCCTTACTC 1013  
QY 916 ACAATGTTGAAATTCGAAACAGAAATGCTCGTATTTATCCCTTCGTTATGTTTCAAA 975  
Db 1014 TCAATATGCTGAAATGGAAAGACGAATCGTGTATTTATTCCTTCGTTTATGTTTCAA 1073  
QY 976 CCATTTGGGTACCAGATTTCAAGACCCAGAAACCAAGTCCCAATCGACTCCGGAACCTAG 1035  
Db 1074 CATTGGGTACAGATTCAGGCGCAAGAAACCAAGTCCCAACCGACTCCGGAACCTAG 1133  
QY 1036 TCCAAAGTCCGCAACCTGCAACCAATCTCAACGAGTCCCAAGCAATCCCAATTCATGAGAA 1095  
Db 1134 TCCAGGCCCGCAACCTGCAACCAATCTTAAATATAGACTCAA-----ATTCTTC 1181  
QY 1096 ATTGTTCAAGAGGCTGTTTCGAAAGTAGCGATGTTGTTATGCTTTGAGGAGATGAGT 1155

1182 TTGGTTAGTCTGCTGGTACGAAAGTTGGGGAAGATATGATTTCGAAGAAAGGCGAT 1241  
1156 TTCTCGTTATATCCAGCAAGATCTTTTCACAGAAAACAGCAGCAGGATGATGACAA 1215  
1242 CTCTCGTTATGTTTGGGAAGATTTACCATCTGAATCTTTAAATCTTTGAAGCAA 1301  
1216 ACTGCCAAGCAGGAAGTTTATCTATAGCTAGAGCTAGAAAACCTGACCTCCATC 1275  
1302 GTTATCAAAAACAGAGAGTGTTCACACACTTTAACTGCTAAAAAGAAAATGTTGCTCC 1361  
1276 TAGTGATCGAGAATTTTCAATTAAGCTTTATGACCTTACGACGAAGATTCACCAAGATT 1335  
1362 TCGTGACCAAGAATTTTATGATAAAGCATATATCTGTTAACTGAGGCTCATAAAGCCTT 1421  
1336 ACTTGATTAAGAGTTCGACAGATTTGATTTTTCAGGCTTTTGGATTAACCTGTGGAAGACT 1395  
1422 GTTTGNAATTAAGGTCGTAATCTGATTTTCCAGCCTTTAGACAAATTAATTAAGACGCTT 1481  
1396 CAAGGATGTCNCAAGTGATAAAGTCAAGTTAGTGGANGATATTTCTTGCTTCTTACTCC 1455  
1482 GAATGATGAATCGACTAATAAGAAAATTTGGTAGATGATTTATTGGCATTTCTAGCACC 1541  
1456 GATTCGTATCCAGAACGTTTATAGAAAACCAAAATGCGCAAAATTAACCTACATGATGATGA 1515  
1542 AATTACCATCCAGAGCGACTTGGCAAAACCAAAATTTCTCAAATTTAGTATATCTGAAGACGA 1601  
1516 GATTCAGATGCAAGTTGGCAGGCAAGTACACAGAGAGCGTTATATCTTTGATCC 1575  
1602 AGTTGCTATGCTCAATGCTGATAAGTATACAGCTCAGATGGTTTACATTTTATGATGA 1661  
1576 TCGTGATATAACACGATGATGAGGGGATCCCTATGTAACCTCCACATATGACCCATAGCCA 1635  
1662 ACATGATATATCACTGATGAGGAGATGATATGTAACGCTCATATGGCCATAGTCA 1721  
1636 CTGGATTAAGAAAAGATGTTTCTGTAAGCTGAGAGCGGCGAGCCGCTTATGCTAA 1695  
1722 CTGGATTCGAAAAGATGAGCTTTCTGATAGGAAAAGTTGAGCTCAAGCCTATATCTAA 1781  
1696 AGAGAAAGTTGACCCCTCTTCGACAGACCATCAGGATTCAGGAATTCAGGAATCTGAGSCAA 1755  
1782 AGAAAAGGTTATCTTCACTCCATCTCCAGACGAGATGTTAAAGCAATCCAACTGGAGA 1841  
1756 AGGACGAGAAGTATCTCAACCGCTGAAAGCAGCTAAGAGGTCCTGATCGTAT 1815  
1842 TAGTGACAGAGCTATTTCAATCTGTTGTAAGGGGAAAACGAAATTCCTGCTGACT 1901  
1816 GCCTTACAACTTCAATATCTGTAAGCTCAAAAACGCTAGTTTAACTCATACCTCATTA 1875  
1902 TCCATATATGTTGAGCATACAGTTGAGGTTAAAAAACGGTAAATTTGATTTATCTCATAA 1961  
1876 TGACCATTAACATAACATAAATTTGAGTTGAGTTGAGGAGGCTTTTGTAGGACCTTAA 1935  
1962 GGATCATTAACATAATATTAATTTGCTTTGTTGATGATCACAATACAAAGCTCCAAA 2021  
1936 GGGGTATCTCTTGAGGATCTTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGA 1995  
2022 TGGCTATACCTTGGAAAGATTTGTTTGGAGCATTAAGTACTAGTAGAACCCCTGACGA 2081  
1996 AGCTCCGATTCAGATAATGTTTGTGTAAGCTAGCGACCATGTTCAAGAAACCAAAA 2055  
2082 AGCTCCACATTTAATGATGATGGGCAATGCCAGTGAGCATGTTGTTAGGCAAGAAAGA 2141  
2056 TGGTCAGCTGATACCAATCAACGGAAGAAAACCAAGCGGAGGAGAAACCTTCAGACAGAAA 2115  
2142 CCACAGTGAAGATCCAAATAAGAACTTCAAGCGGATGAGAGCCAGTAGAGGAAACACC 2201  
2116 ACCTGAGGAGAAACCCCTCGAAGAGA 2144  
2202 TGCTGAGCCAGAAGTCCCTCAAGTAGAGA 2230

RESULT 16

US-10-387-783-9

; Sequence 9, Application US/10387783  
; Publication No. US2004000531A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; TITLE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-683  
; CURRENT APPLICATION NUMBER: US/10/387,783  
; CURRENT FILING DATE: 2003-03-13  
; PRIOR APPLICATION NUMBER: 09/468,656  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 9  
; LENGTH: 2451  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(2451)  
; OTHER INFORMATION: n = a, c, t or g  
US-10-387-783-9

Query Match 43.38; Score 990.6; DB 16; Length 2451;  
Best Local Similarity 67.84; Pred. No. 1.3e-267;  
Matches 1484; Conservative 0; Mismatches 645; Indels 60; Gaps 5;

QY 1 TTGTTCTCTAAGAACTCGTCTGCTACCAAGCTGCTCAGGTTAAGAAAAGAGTCTAAATCGAGT 60  
Db 57 TTGTTCTTACAGTTTGGGACTGTATCAAGCTAGAACGGTTAAGGAAA---TAATCGTGT 113  
QY 61 TTCTTATATAGATGGTGATCAGGCTGGTCAAAAGCGAGAAAACCTTGACACAGATGAAGT 120  
Db 114 TTCTTATATAGATGGAATAACAAGCGACGCAAAAACCGAGAAATTTGACTCTCTGATGAGT 173  
QY 121 CAGTAAGAGGAGGGGATCAACGCCGAACAAATGTTATCAAGATTACGATTCAGGATCAAGTTA 180  
Db 174 TAGCAAGCGTGAAGGAATCAATGCTGAGCAATCGTCAATCAAGATAACAGACCAAGGCTA 233  
QY 181 TGTGACCTCTCATGCGAGACCATTAATCAATTAATGCGCAAGGTTCTTATGATGCCAT 240  
Db 234 TGTCACTTCATCGCGACCACTATCAATTAATCAATGTAAGGTTCTTATGACGCTAT 293  
QY 241 CATCAGTGAAGAGCTCCTCATGAAGATCCGAATTTATCAGTTTGAAGGATTCAGACATTTGT 300  
Db 294 CATCAGTGAAGAAATTAATCATGAAGATCCAACTATAAGCTAAAAGATGAGGATATTGT 353  
QY 301 CAATGAATCAAGGGTGGTTATGTCATTAAGGTAAACGGTAAATCAATGTTACCTTAA 360  
Db 354 TAATGAGGTCAAGGGTGGATATGTTATCAAGGTAGATGGAATACTATGTTTACTTTAA 413  
QY 361 GGATCGAGCTCATCGGATTAATTCGACAAAAGAGAGATTAACCTGCAAGACAGA 420  
Db 414 GGATGCTGCCACCGGATTAACGTCGTAACAAAAGAGAAATCAATCGACAAAACAAGA 473  
QY 421 ACGCAGTCATAATCAT-----AACTCAAGAGCAGATAATGCTGTTGCTGACGCCAG 471  
Db 474 GCATAGTCAACATCGTGAAGTGGAACTCCAAAGAAAACGATGGTGTGTTGCTTGGCAG 533  
QY 472 AGCCCAAGGAGCTTATCAACGGATGATGGGTATATCTTCAATGCAATCATGATCATTTGA 531  
Db 534 TTCGAAGGAGCTTATCAATCAGATGATGTTATATCTTTAATGCTTCTGATATCATAGA 593  
QY 532 GGACACGGGTGATCTTATATCGTTCTTCAAGGACCATTAACCATTAATTCCTTAAGAA 591  
Db 594 GGATACCTGGTGAATCTTATATCGTTCTCATGGAGATCATTAACCATTAATTCCTTAAGAA 653  
QY 592 TGAGTTATCAGTAGCGAGTTAGTCTGCTGCAAGACCT----- 629

Db 654 TCAGTTATCAGCTAGCGAGTTGGCTGTGTCGAGAAGCCCTTCCTATCTGCTGCGAGAAATCT 713  
QY 630 -----ATTGGAATCGGAAGCAGGATCTGCTCTTCAAGTTCTAGTTATAA 678  
Db 714 GTCAAAATTCAGAACCTATCGCGCAAAATAGGATACACTTCAGAAACAACCTGGT 773  
QY 679 TGCAAAATCCAGCTCAACCAAGATTTGTCAGAGAACCACAAATCTGACTGCTCCTCAACTTA 738  
Db 774 ACCTTCTGTAAAGCAATCCAGAACTCAAAATACTAACACAGCAACAACAGCACTAA 833  
QY 739 TCATCA---AAATCAGGGGAAACAATTCAGCCCTTTTAGGTAATGATGCTAAACC 795  
Db 834 CAGTCAAGCAAGTCAAAAGTAATGACATTTGATAGTCTCTTGAACAGCTCTACAAACTGCC 893  
QY 796 CTTATCAGACGCCATGTCGAATCTGATGCTTATTTTCAGCCAGCGCAAAATCACAAG 855  
Db 894 TTTGAGTCAACGACATGTAGAATCTGAATGCTGCTCTTTGATCCAGCAAAATCACAAG 953  
QY 856 TCGAACCGCCAGAGGTAGTCTGCTCCTCATGTGTAACCAATTTATCCCTTATGA 915  
Db 954 TCGAACAGCTAGAGGTGTCAGTGCACACAGGATCAATACCACTTCATCCCTTACTC 1013  
QY 916 ACAATGCTGTAATGGAAAAAGAAATGCTGTAATTTCCCTCTGTTATGCTTAA 975  
Db 1014 TCAATGCTGTAATTGAAGAAAGAAATGCTGTAATTTCCCTCTGTTATGCTTCAAA 1073  
QY 976 CAAATGGGTACAGATTCAGAACAGAACCAAGTCCCAATTCGACTCCGGAACCTAG 1035  
Db 1074 CCATGGGTACCAAGATTCAGAGCCAGAACCAAGTCCCAACCGACTCCGGAACCTAG 1133  
QY 1036 TCCAAGTCCGCAACTCCACCAATCTCAACAGCTCCAGCAATCAATTTGATGAA 1095  
Db 1134 TCAGGCCCGCAACTCCACCAATCTCAATATGACTCA-----ATTCTTC 1181  
QY 1096 ATTGGTCAAAAGAGCTTTCGAAAGTAGGCGATGTTATGCTTTGAGGAGATGGAGT 1155  
Db 1182 TTTGGTTAGTCAGCTGTACGAAAGTTGGGAAGGATATGTTTGAAGAAAGGCGAT 1241  
QY 1156 TTTCTGTTATATCCAGCAGGATCTTTGAGAGAAACAGCAGCAGGCAATGTAGCAA 1215  
Db 1242 CTCTCGTTATGCTTTGCGAAGATTTACCATCTGAAACTTTAAATCTTTGAAGCAA 1301  
QY 1216 ACTGGCCAGCAGAAAGTTTATCTATAAGCTAGGAGTAAAGAACTGACCTCCCATC 1275  
Db 1302 GTTATCAAAACAGAGAGTGTTCACACACTTTACTGCTAAAAGAAATGTTGCTCC 1361  
QY 1276 TAGTGATPCGAGAAATTTTACAATAAGGCTTATGACTTACTAGCAAGAAATTCACCAAGATT 1335  
Db 1362 TCGTGACCAAGAAATTTATGATAAGCATATAATCTGTTAACTGAGGCTCATAAAGCCTT 1421  
QY 1336 ACTTGATAATAAGGTGCAACAGTTGATTTGAGGCTTTGATTAACCTGTTGGAACGACT 1395  
Db 1422 GTTTGNAATAAAGGCTGTAACTGATTTTCAAGCCCTTAGCAAAATTTAGAACGCTT 1481  
QY 1396 CAGGATGTCNAGTGTAAAGTCAAGTTAGTGGAGATATCTTGCTCTTTAGCTCC 1455  
Db 1482 GAATGATGAATCGACTAATAAGAAATTTGGTAGATGATTTATGGCATTTCTTAGCACC 1541  
QY 1456 GATTCGTCATCCAGAGCTTTAGGAAACCAATTCGCAAAATTTACTACACTGATGATGA 1515  
Db 1542 AATTAACCAATCCAGGCGACTTGGCAACCAATTTCTCAATTTGATATACTGAAGACGA 1601  
QY 1516 GATTCAGATAGCAGTTGGCAGCAAGTACACAAAGAGACGGTTATATCTTTGATGCC 1575  
Db 1602 AGTTGCTGATTTCTCAATAGCTGATAAGTATACACAGTCAAGATGTTTATGATGA 1661  
QY 1576 TCGTGATATAACAGTGTAGGGGATGCTTACTTCAATCCACATATGACCATAGCCA 1635  
Db 1662 ACATGATATAATCAGTGTAGGAGATGCTATGTAACGCTCATATGGGCCATAGTCA 1721  
QY 1636 CTGATTAATAAAGAGATGTTTGTCTGAAGCTGAGAGAGCGGCGAGCCAGGCTTATGCTAA 1695  
Db 1722 CTGATTTGAAAGAGATGCTTTCTGATAAGAAAGTTTGCAGCTCAAGCCTACTATA 1781

QY 1696 AGAGAAAGTTTGGACCCCTCTCTCGACAGACCATCAGGATTCAGAAATCTGAGGCAAA 1755  
Db 1782 AGAAAAAGTATCTACCTCTCATCTCCAGACGCGAGATGTAAAGCAAAATCCAACCTGGAGA 1841  
QY 1756 AGGAGCAGAACCTATCTACAACCGGCTGAAGCAGCTGAAGAGGTGCCACTTGTATCGTAT 1815  
Db 1842 TAGTGCAGCAGCTATTTACAATCGTGTGAAAAGGGGAAAAACGAATTCCACTCGTTCGACT 1901  
QY 1816 GCCTTACAATCTCAATATACGTAGAAAGTCAAAAACGGTAGTCTTTAATCATACCTCATTA 1875  
Db 1902 TCCATATATGTTGAGCATACAGTTGAGTTTAAACCGTAAATTTGATTTCTCTCATAA 1961  
QY 1876 TGACCAATTACCATAACCAATTTGAGTGGTTTGAACGAGCCCTTTATGAGGACCTTA 1935  
Db 1962 GGATCAATTACCATAATATAAATTTGCTTGTGTTGATGATCACACATACAAAGCTCCAA 2021  
QY 1936 GGGGTATATCTCTGAGGATCTTTTGGCGACTCTCAAGTACTATGTCAGACATCCAAACGA 1995  
Db 2022 TGGCTATATCTTGGAGATTTGTTTTCGACGATTAAGTACTAGTAGAACAACCCCTGACGA 2081  
QY 1996 ACCTCGCATTCAGATAATGTTTGTAACTGCTAGCGACCATGTTCAAAGAAACAAAA 2055  
Db 2082 ACCTCCACATTTTAATGATGATGGGCAATGCCAGTGAAGTGTAGGCAAGAAAGA 2141  
QY 2056 TGGTCAAGCTGATACCAATCAAAACGAAAAAACCAAGCGAGGAGAAACCTCAGACAGAAA 2115  
Db 2142 CCACAGTGAAGATCCAAATAAGAACTTCAAAGCGGATGAAGACCGAGTAGAGAAACACC 2201  
QY 2116 ACCTGAGGAAGAAACCCCTCGAGAGAGA 2144  
Db 2202 TGCTGAGCGAAGATCCCTCAAGTAGAGA 2230

## RESULT 17

US-09-765-272-55  
; Sequence 55, Application US/09765272  
; Patent No. US20020061545A1  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/765,272  
; FILING DATE: 22-Jan-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 55:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2389 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 55:

S-09-765-272-55

Query Match 43.1%; Score 987.6; DB 9; Length 2389;  
Best Local Similarity 67.7%; Pred. No. 9.2e-267;  
Matches 1481; Conservative 0; Mismatches 645; Indels 60; Gaps 5;

4 TTCCTATGAACTCGTCTGCTACCAAGCTGGTTCAGGTTAAGAAAGAGTCTTAATCGAGTTTC 63  
1 TTCTTACAGTTGGAGTGTATCAAGCTAGAACGTTAAGGAAA--TAAICGTGTTTC 57  
64 TTATATAGATGGTATCAGCTGGTTCAAAAGCAGAAAACCTTGAACACAGATGAAGTCAG 123  
58 CTAATATAGTGAAGAAACAAGCAGCAAAAACCGAGAAATTTGACTCTCTGATGAGGTTAG 117  
124 TAAGAGGAGGGGATCAACGCGGAACAATNGTNAATCAAGATTACGGATCAAGTTATGT 183  
118 CAAGCGTGAAGGAATCAATGCTGAGCAATTCGTCAAGATAACAGACCAAGGCTATGT 177  
184 GACCTCTCATGAGACCAATATCATTAATGCGCAAGTTTCTTATGATGCCATCAT 243  
178 CACTTCAATGCGGACCACTATCATTAATTAACAATGTAAGGTTCTTATGACGCTATCAT 237  
244 CAGTGAAGAGCTCCTCATGAAGATCCGAAATTTATCAGTTGAAGGATTCAGACATTCCTAA 303  
238 CAGTGAAGATTTACTCATGAAGATCCAAACTATTAAGCTAAAGATGAGATATTTGTTAA 297  
304 TGAATCAAGGGTGGTATGTGATTAAGTTAAAGGTAATTAATTAATTAATTAATTAATTA 363  
298 TGAGTCAAGGGTGGTATGTGATTAAGTTAAAGGTAATTAATTAATTAATTAATTAATTA 357  
364 TGCAGCTCATGCGGATTAATTTCCGACAAAGAGAGATTAACAGTCAGAAAGCAGAAAG 423  
358 TGCTGCCACGCGGATTAAGTCCGTAACAAAGAGGAAATCAATCGACAAACAAAGACGA 417  
424 CAGTCATATATCAT-----AACTCAAGAGCAGATAATGTCTGTCTGTCAGCCAGAGC 474  
418 TAGTCAACATCTGTAAGTGGAACTCCAGAAACGATGGTGTCTGTTGCCCTTGCAAGTTC 477  
475 CCAAGGAGCTTTATACAACGGATGATGGGTATATCTTCAATGGAATCTGATATCATTAAGGA 534  
478 GCAAGGAGCTTATATACAGATGATGGTATATCTTTAATGCTTCTGATATCATAGAGGA 537  
535 CACGGTGTATGCTTATATGTTCTCAGCGGACCATTAACATTAATTCCTTAAGATGA 594  
538 TACTGTGTATGCTTATATGTTCTCATGGAGATCATTAACATTAACATTCCTAAGATGA 597  
595 GTTATCAGCTAGCGAGTTAGTCTGTCAGAGGCT----- 629  
598 GTTATCAGCTAGCGAGTTGGTCTGTCAGAAAGCTTCTATCTGGTCGAGGAATCTGTC 657  
630 -----ATTGGAATGGGAAGCAGGAGATCTCGTCTTCTCAAGTTCTAGTTATTAATGC 681  
658 AAATTCAGAAACCTATCGCCGCAAAAATACGATATACACTTCAAGAACAAACTGGGTACC 717  
682 AAATCCAGCTCAACCAAGATGTGAGAGAACCAATCTGACTGTCACTCCAACTTATCA 741  
718 TTCTGTAAGCAATCCAGGAATCAAAATACTAAACAAGCAACAAACAGCAACACTAACAG 777  
742 TCA--AAATCAAGGGGAAACATTTCAAGCTTTTACGTGAATTTGATGTAAACCTTT 798  
778 TCAAGCAAGTCAAGATTAATGACATTTGATGTCTTTGAAACAGCTCTACAACTGGCCTTT 837  
799 ATCAGAACGCAATGTGGAATCTGATGCGCTTATTTTCGACCCAGCGCAAAATCAAGTCG 858  
838 GAGTCAACGACATGTAGAACTGATGCGCTTGTCTTTGATCCAGCAAAATCAAGTCG 897  
859 AACCGCCAGAGGTGTAGCTGTCCCTCATGGAATTAACATTAACATTTATCCCTTATGAACA 918  
898 AACAGCTAGAGGTGTTGCAAGTGCCACACGAGAGATCAATACCACTTCTATCCCTTCTCA 957  
919 AATGCTGAATTTGAAAAACGAATGCTCGTATTTATTTCCCTTCGTTATCGTTCAAACCA 978

Db 958 AATGCTGAATTTGGAAGACGATCGCTCGTATTTATTTCCCTTCGTTATCGTTCAAACCA 1017  
QY 979 TTGGGTAC CAGATTCAAGACCAGAA CAACAAGTCCCAATTCGACTCGGAACTAGTCC 1038  
Db 1018 TTGGGTAC CAGATTCAAGGCCAGAA CAACAAGTCCCAACCGGACTCGGAACTAGTCC 1077  
QY 1039 AAGTCCGCAACCTGCAACCAATCTCTCAACAGCTCCAAAGCAATCCAAATGATGAAAT 1098  
Db 1078 AGCCCGCAACCTGCAACCAATCTCTCAACAGCTCCAA--ATTCTCTTT 1125  
QY 1099 GGTCAAGAGCTGTTTCGAAAGTAGCGATGTTATGTCTTTGAGGAGATGAGTTTC 1158  
Db 1126 GGTATGCTGAGTGTGAGAAAGTTGGGAGAGGATATGTTTCGAGAAAGGCACTC 1185  
QY 1159 TCTTATATCCAGCCCAAGGATCTTTTCAGCAGAAACAGCAGCGCATTTGATAGCAACT 1218  
Db 1186 TCTTATGCTTTGCGAAGATTTACCATCTGAAACTGTTAAAAATCTTGAAGCAAGTT 1245  
QY 1219 GGCAGAGCAGAAAGTTTATCTCAATAGCTAGGAGCTAAGAAACTGACCTCCCATCTAG 1278  
Db 1246 ATCAAAACAAGAGAGTGTTCACACACTTTAACTGCTAAAAAAGAAATGTTGCTCCCTG 1305  
QY 1279 TGATCGAATTTTACAAATAGGCTTATGCTTACTAGCAAGATTTACCAAGATTTACT 1338  
Db 1306 TGACCAAGAAATTTATGATAAGCATATAATCTGTTAACTGAGGCTCATTAAGCCTGTT 1365  
QY 1339 TGATAATAAAGGTCGACAAAGTTGATTTTGGAGCTTTGGATTAACCTTTGGAAACGACTCA 1398  
Db 1366 TGNAAATAGGGTCTGTAATCTGATTTCCAAAGCTTTAGCAAAATTTATTAAGCCTTGAA 1425  
QY 1399 GGATGTCNCAAGTGAATAAGTCAAGTTAGTGGAGATATCTTGGCTTCTTAGCTCCGAT 1458  
Db 1426 TGATGAATCGACTAATAAAGAAAAATTTGGTAGATGATTTATTTGGCATTTCTTAGCACT 1485  
QY 1459 TCGTCAATCCAGAAAGTTTAGAAACCAAAATTCGCAAAATTTACTACATGATGATCAGAT 1518  
Db 1486 TACCATTCAGAGCGACTTGGCAACCAAAATTTCTCAAATTTAGTATATCTGAAGACGAGT 1545  
QY 1519 TCAAGTAGCCAAAGTTGGCAGGCAAGTACACAAAGAGACGTTTATCTTTGATCCTCG 1578  
Db 1546 TCGATTTGCTCAATTAGCTGATAGTATACAAAGCTGATGATGTTTATGATGAACA 1605  
QY 1579 TGATATAACCAAGTGAAGGGGATGCTATGTAATCTCCACATATGACCCATAGCCACTG 1638  
Db 1606 TGATATAATCAGTGAAGGAGATGCAATGTAACGCTCATATGGGCCATAGTCACTG 1665  
QY 1639 GATTAAAAAAGATAGTTTGTCTGAGCTGAGAGAGCGGCGGCGCTTATGCTAAAGA 1698  
Db 1666 GATTGGAAGATAGCCTTTCTGATAAGGAAAGTTGAGCTCAAGCTTATCTAAAGA 1725  
QY 1699 GAAAGTTTGACCCCTCTTCGACAGACCATCAGGATTCAGGAAATPACTGAGGCAAAAGG 1758  
Db 1726 AAAAGTATCTTACTCTCATCTCCAGAGCGAGATGTTAAAGCAAAATCCAACTGGAGATAG 1785  
QY 1759 AGCAGAGCTATCTACAAACCGCTGAAAGCAGCTAAGAGGTCGCACTTTGATGTATGCC 1818  
Db 1786 TGCAGCAGCTATTTACAATCGTGTGAAAGGGGAAAAACGAAATTCCTACTCGTCTGCTCC 1845  
QY 1819 TTAACAATCTCAATATATCTGTAGAAAGTCAAAACCGTAGTTTAAATCATACCTCATATGA 1878  
Db 1846 ATATATGTTGAGCATACAGTTGAGGTTAAAAACCGTAAATTTGATTTATTTCTCTCATAGGA 1905  
QY 1879 CCATTACCATCAATCAAAATTTGAGTGGTTTTCAGAAAGGCTTTTATGAGCACTTAAGGG 1938  
Db 1906 TCATTACCATCAATTAATAATTTGCTTTGTTTGTATGATCAACATCAAAAGCTCAATGG 1965  
QY 1939 GTATATCTCTGAGGATCTTTTGGCGACTGCTCAAGTACTATGTGCAACATCCAAACGAAAG 1998  
Db 1966 CTATACCTTGGAGATTTGTTTGGACGATTAAGTACTAGTAGAACACCTCTGACGAAAG 2025  
QY 1999 TCCGCAATTCAGATAATGTTTGGTAACGCTAGCGACCATGTTTCAAGAAAGAAACAAATGG 2058

Db 2026 TCACATCTTAAATGATGGGCAATGCCAGTGAGCATGTGTTAGCGAAGAAAGACCA 2085  
QY 2059 TCAAGCTGATACCAATCAACGGAAGAAACCAAGCGAGGAGAAACCTCAGACAGAAAAC 2118  
Db 2086 CAGTGAAGATCCAAATAAGAACTTCAAGCGGATGAAGCCAGTAGAGGAAACACCTGC 2145  
QY 2119 TGAGGAAGAAACCCCTCGAAGAGA 2144  
Db 2146 TGAAGCCAGAGTCCCTCAAGTAGAGA 2171

RESULT 18  
US-10-158-844-355  
; Sequence 355, Application US/10158844  
; Publication No. US20040029118A1  
; GENERAL INFORMATION:  
; APPLICANT: Kunsch et al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-R  
; COMPUTER: Dell Latitude Pentium 3  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/158,844  
; FILING DATE: 03-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/961,527  
; FILING DATE: 1997-10-30  
; APPLICATION NUMBER: US 60/029,960  
; FILING DATE: 1996-10-31  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hyman, Mark J.  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB340PLD1  
; INFORMATION FOR SEQ ID NO: 355:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 973 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 355:  
US-10-158-844-355

Query Match 34.7%; Score 794.8; DB 13; Length 973;  
Best Local Similarity 98.7%; Pred. No. 1e-212;  
Matches 810; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 1 TTGTTCTTATGAACCTTGGTCGTCACCAAGCTGTCAGGTTAAGAAAGAGTCTTAATCAGT 60  
Db 154 TTGTTCTTATGAGCTTGGACGTTACCAAGCTGTCAGGATGAAGAAAGAGTCTTAATCAGT 213  
QY 61 TTCCTTATATAGATGGTATCAGGCTGGTCGAAGGCGAAGAACTTCACACAGATGAAGT 120  
Db 214 TGCCTTATATAGATGGTATCAGGCTGGTCGAAGGCGAAGAACTTCACACAGATGAAGT 273  
QY 121 CAGTAAGAGGAGGAGGATCAACCCGCAACAAATGATNATCAAGATTACGGATCAAGTTA 180  
Db 274 CAGTAAGAGGAGGAGGATCAACCCGCAACAAATGATNATCAAGATTACGGATCAAGTTA 333  
QY 181 TGTGACCTCTCATGGAGACCATTAATCATTAATTAATGGAAGGTTCTTATGATGCCAT 240  
Db 334 TGTGACCTCTCATGGAGACCATTAATCATTAATTAATGGAAGGTTCTTATGATGCCAT 393

QY 241 CATCAGTGAAGAGCTCCTCATGAAGATCCGAATTAATCAGTTGAAGGATTCAGACATTTGT 300  
Db 394 CATCAGTGAAGAGCTCCTCATGAAGATCCGAATTAATCAGTTGAAGGATTCAGACATTTGT 453  
QY 301 CAATGAATCAAGGCTGTTATGTCATTAAGGTAACGTAATTAATTAATTAATTAATTAAT 360  
Db 454 CAATGAATCAAGGCTGTTATGTCATTAAGGTAACGTAATTAATTAATTAATTAATTAAT 513  
QY 361 GGATGCAGCTCATGGGATTAATTTCCGACAAAAGAGATTTAAACGTCAGAAAGCAGGA 420  
Db 514 GGATGCAGCTCATGGGATTAATTTCCGACAAAAGAGATTTAAACGTCAGAAAGCAGGA 573  
QY 421 ACGCAGTCATTAATCAATACTCAAGAGCAGATTAATGCTGTGTCAGCCAGAGCCCAAGG 480  
Db 574 ACGCAGTCATTAATCAATACTCAAGAGCAGATTAATGCTGTGTCAGCCAGAGCCCAAGG 633  
QY 481 ACGTTATACAACGGATGATGGGTATATCTTCAATGATCTGATATCATTTGAGGACACGGG 540  
Db 634 ACGTTATACAACGGATGATGGGTATATCTTCAATGATCTGATATCATTTGAGGACACGGG 693  
QY 541 TGATGCTTATATGCTTCTCAGCGGACCATTTACCATTACATTCCTAAAGAAATGAGTTATC 600  
Db 694 TGATGCTTATATGCTTCTCAGCGGACCATTTACCATTACATTCCTAAAGAAATGAGTTATC 753  
QY 601 AGCTAGCGAGTTAGCTGCTGCAAGAGCCCTATTGGAATGGAAAGCAGGATCTCGTCTCTTC 660  
Db 754 AGCTAGCGAGTTAGCTGCTGCAAGAGCCCTATTGGAATGGAAAGCAGGATCTCGTCTCTTC 813  
QY 661 TTCAAGTTCTAGTTATTAATGCAAAATCCAGCTCAACCAAGATTCAGAGAAACCAATCT 720  
Db 814 TTCAAGTTCTAGTTATTAATGCAAAATCCAGCTCAACCAAGATTCAGAGAAACCAATCT 873  
QY 721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGA 780  
Db 874 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGA 933  
QY 781 ATTGTATGCTAAACCTTATCAGAACGCCATTTGGAATCTG 821  
Db 934 ATTGTATGCTAAACCTTATCAGAACGCCATTTGGAATCTG 973

RESULT 19  
US-10-158-844-258  
; Sequence 258, Application US/10158844  
; Publication No. US20040029118A1  
; GENERAL INFORMATION:  
; APPLICANT: Kunsch et al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-R  
; COMPUTER: Dell Latitude Pentium 3  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/158,844  
; FILING DATE: 03-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/961,527  
; FILING DATE: 1997-10-30  
; APPLICATION NUMBER: US 60/029,960  
; FILING DATE: 1996-10-31  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hyman, Mark J.  
; REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB340PID1

INFORMATION FOR SEQ ID NO: 258:

## SEQUENCE CHARACTERISTICS:

LENGTH: 1684 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 258:

US-10-158-844-258

Query Match 29.6%; Score 677; DB 13; Length 1684;  
Best Local Similarity 99.9%; Pred. No. 2.3e-179;  
Matches 688; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1602 ATGCTATGTAATCACTACATATGACCATAGCCACTGGATTAAAGAGATAGTTGCTG 1661  
DB 1 ATGCTATGTAATCACTACATATGACCATAGCCACTGGATTAAAGAGATAGTTGCTG 60

QY 1662 AAGCTGAGAGCGGCGAGCCAGGCTTATGCTAAAGAGAAAGGTTTGACCCCTCCTTCGA 1721  
DB 61 AAGCTGAGAGCGGCGA-CCAGGCTTATGCTAAAGAGAAAGGTTTGACCCCTCCTTCGA 119

QY 1722 CAGACCATCAGGATTCAGGAATATCTAGGCAAAAGGAGCAGAGTATCTCAACGCG 1781  
DB 120 CAGACCATCAGGATTCAGGAATATCTAGGCAAAAGGAGCAGAGTATCTCAACGCG 179

QY 1782 TGAAGCAGCTAAGAGGTGCGAATCTGATGCTTACATCTTCAATATCTATCTAG 1841  
DB 180 TGAAGCAGCTAAGAGGTGCGAATCTGATGCTTACATCTTCAATATCTATCTAG 239

QY 1842 AAGTCAAAACGCTAGTTTAAATCATACCTCATATGACCAATACCAATCAATTTG 1901  
DB 240 AAGTCAAAACGCTAGTTTAAATCATACCTCATATGACCAATACCAATCAATTTG 299

QY 1902 AGTGGTTTCAAGAGCCCTTTATGAGGCACTAAGGGGTATCTCTTGAGGATCTTTTG 1961  
DB 300 AGTGGTTTCAAGAGCCCTTTATGAGGCACTAAGGGGTATCTCTTGAGGATCTTTTG 359

QY 1962 CGACTGTAAGTACTATGCGAATCATCCAAACGAAAGTTCGCAATTCAGATAGTTTG 2021  
DB 360 CGACTGTAAGTACTATGCGAATCATCCAAACGAAAGTTCGCAATTCAGATAGTTTG 419

QY 2022 GTAAAGCTAGGACCATGTTCAAGAAACAAATGCTCAAGCTGATACCAATCAACGG 2081  
DB 420 GTAAAGCTAGGACCATGTTCAAGAAACAAATGCTCAAGCTGATACCAATCAACGG 479

QY 2082 AAAAACCGAGGAGAGAAACCTCAGACAGAAACCTGAGGAGAAACCCCTCGAGAAG 2141  
DB 480 AAAAACCGAGGAGAGAAACCTCAGACAGAAACCTGAGGAGAAACCCCTCGAGAAG 539

QY 2142 AGAAACCGAAAGCGAGAAACCGAGTCTCCAAACCAACAGAGGAAACCGAGAAATCAC 2201  
DB 540 AGAAACCGAAAGCGAGAAACCGAGTCTCCAAACCAACAGAGGAAACCGAGAAATCAC 599

QY 2202 CAGAGAAATCAGAGAACCTCAGCTCGAGCTGAAAGGTTGAGAGAAACCTGAGAGAG 2261  
DB 600 CAGAGAAATCAGAGAACCTCAGCTCGAGCTGAAAGGTTGAGAGAAACCTGAGAGAG 659

QY 2262 CTGAGATTTACTTGGAAAAATCCAGAT 2290  
DB 660 CTGAGATTTACTTGGAAAAATCCAGAT 688

RESULT 20

US-09-765-272-181

; Sequence 181, Application US/09765272

; Patent No. US20020061545A1

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,272  
FILING DATE: 22-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/961,083  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 181:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1342 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 181:  
US-09-765-272-181

Query Match 17.5%; Score 401; DB 9; Length 1342;

Best Local Similarity 65.7%; Pred. No. 1.1e-101;

Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

QY 23 CACCAAGCTGCTCAGGTTAAGAAAGAGCTAATCGAGTTCTTATATAGATGGTATCAG 82  
DB 8 CAGATCGTTCGAGGAAAATAAGCAATATCGTGTCTTATGTGATGGCAGCCAG 67

QY 83 GCTGCTCAAAAGGCGAGAAACCTTGACACAGATGAAGTCAGTAAGAGGGAGGATCAAC 142  
DB 68 TCAAGTCAGAAAGTGAAACCTTGACACAGACAGGTTAGCCAGAAAGAAATTCAG 127

QY 143 GCCGAAACAAATNGVATCAAGATTCAGGATCAAGTTATGTGACCTCTCATGGAGCCAT 202  
DB 128 GCTGAGCAAAATTTGAATCAAAATTTACAGATCAGGCTATGTAACGTCACCGTGACCAC 187

QY 203 TATCATTAATAATGCGCAAGGTTCTTATGATGCCATCATCAGTGAAGAGCTCCTCATG 262  
DB 188 TATCATTAATAATGCGAAAGTTCTTATGATGCCCTTTAGTGAAGAACTTTGATG 247

QY 263 AAAGATCCAAATTAATCAAGTGAAGGATTCAGACATTTGCAATGAATCAAGGGTGGTAT 322  
DB 248 AAGATCCAAATTAATCAAGTGAAGGATTCAGACATTTGCAATGAATCAAGGGTGGTAT 307

QY 323 GTCATTAAGGTAACCGTAAATTAATCTATGTTACCTTAAGGATGCAGCTCATGGGTAAT 382  
DB 308 ATCATCAAGGTCGATGGAATAATTTATGTTCTACCTGAAAGATGCAGCTCATGCTGAT 367

QY 383 ATTCGACAAAGAGAGATTAAGCTCAGAGCGAGGACCGAGTCATATCAATCAATCA 442  
DB 368 GTTCGAACTAAAGATGAATCAATCGTCAAAACAAAGAAC---ATGCAAGTAATAG 424

QY 443 AGAGCAGATAATGCTGTTGCTGACGAGCCAGAGCCGAGGAGGAGGAGGAGGAGGAG 502  
DB 425 AAGGTTAACTTAATGTTGCTGTAGCAAGGTTCTCAGGAGGAGGAGGAGGAGGAGG 484

QY 503 TATATCTTCAATGCAATCTGATATCAATTTAGGACACGGGGTGGTATATCGTTCCTCAC 562  
DB 485 TATGCTTTAATCCAGCTGATATTTATCGAAGATACGGGTAATGCTTATATCGTTCCTCAT 544



563 GGCACCATTTACCAATCAATCTCTAAGAAATGAGTATATACAGCTAGCGAGTTAGTCTGCGA 622  
545 GGAGTCTACTATCACTACATCTCCCAAGAGCGATTATCTGCTAGTGAATAGCAGCAGCT 604  
623 GAAGCTATTGGAATGGGAAGCAGGAGTCTCGTCTCTTCTCAAGTTCTAGTTATATGCA 682  
605 AAAGCACATCTGGCTGGAAAAAATATGCAACCGAGTCAGTTAGCTATTCTTCAACAGCT 664  
683 AATCAGCTCAACCAAGATTGTCTGAGAGAACCAATCTGACTCTCTCACTCAACTTATCAT 742  
665 AGTGACAAATAACCGCAATCTGTAGCAAA-----AGGATCAACTAGCAGCCA 712  
743 CAAATCAAGGGGAAACAAATTTCAAGCTTTTCAAGTGAATTTGATGCTAATAACCTTATCA 802  
713 GCATATAATCTGAAATCTCCAGAGTCTTTGAAGGAATCTTATGATTCACCTAGCGCC 772  
803 GAACGCATGTGGAATCTGATGGCTTTATTTTCGACCCAGCGCAATATCAAGTCGAAAC 862  
773 CAACGTTACAGTGAATCAGATGGCTGGTCTTTGACCTGCTAAGATTATCACTCGTACA 832  
863 GCAGAGGTGTAGTCTCGCTCATGTAAACCATTAACCTTATCCCTTATGAACAAATG 922  
833 CCAAAATGGAGTTGCGATTCGGCATTCGGCATTCGGCAATTAACCACTTTATTCCTTACGCAAGCTT 892  
923 TCTGAATTCGAAAAACGAATTTGCTGCTATTATTTCCTTCCT 964  
893 TCTGCTTAGAAGAAAGATTGCGAATGGTGGCTATCACT 934

RESULT 21  
US-09-769-787-246  
; Sequence 246, Application US/09769787  
; Publication No. US20030091577A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbial  
; APPLICANT: Gilbert, Christophe  
; APPLICANT: Hansbro, Philip M  
; TITLE OF INVENTION: Proteins  
; FILE REFERENCE: PWC/P21129WO  
; CURRENT APPLICATION NUMBER: US/09/769,787  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: GB 9816337.1  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: US 60/125164  
; PRIOR FILING DATE: 1999-03-19  
; NUMBER OF SEQ ID NOS: 388  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 246  
; LENGTH: 1455  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-09-769-787-246  
Query Match 17.5%; Score 401; DB 10; Length 1455;  
Best Local Similarity 65.7%; Pred. No. 1.1e-101;  
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;  
23 CACCAAGCTGCTCAGGTTAAGAAAGAGTCTAATCAGTTCTTATATAGATGGTATCAG 82  
79 CAGCATGTTCCGAGGAAATAAGCAATAATCTGCTCTTATGTGGATGGCAGCCAG 138  
83 GCTGGTCAAAAGGCAAAATTTGACACCATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 142  
139 TCAAGTCAGAAAGTGAATTTGACACCAAGGTTAGCCAGCAAGGAAAGAAATTCAG 198  
143 GCCCAAGAAATNGTATCAAGATTACGATCAAGTATGATGCTATGCTATGATGATGATGATGAT 202  
199 GCTGAGCAATTTGATCAAAATTAAGATCAGGCTATGATGATGATGATGATGATGATGATGAT 258  
203 TATCATTTACTAATGGAAGGTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 262  
259 TATCATTTACTAATGGAAGGTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 318

263 AAAGATCCGAATTTATCAGTTGAAGGATTCAGACATTTGTCATGAAATCAAGGTTGTTAT 322  
319 AAGGATCCAAACTATCAACTTAAAGACGCTGATATTGTCAATGAAGTCAAGGTTGTTAT 378  
323 GTCAATTAAGGTAAACGGTAAATACTATGTTTACCTTAAGGATGAGGATTCATGCGGATAAT 382  
379 ATCATCAAGGTGATGGAATATTTATGCTACCTGAAAGATGAGGATTCATGCTGATAT 438  
383 ATTGCGCAAAAGAGAGATTAAACGTCAGAGGAGGAGGAGGATTCATATTAATCAATCA 442  
439 GTTCGAATTAAGATGAATCAATCGTCAAAACAAGAAC-----ATGTCAAAAGATTAATGAG 495  
443 AGAGCAGATTAATGCTGTTGCTGTCAGGACGAGCCAGGAGCTTATACACGATCATGGG 502  
496 AAGGTTAACTCTAATGTTGCTGTAGCAAGGTTCTCAGGACGATATACGCAATTAATGAT 555  
503 TATATCTTCAATGATCTGATATCAATGAGGACAGCGGTGATGCTTATATCGTTCTCTCAC 562  
556 TATGTCCTTTAATCCAGCTGATTTATCGAAGATACGGGTAAATGCTTATATCGTTCTCAT 615  
563 GGCACCATTTACCAATTACCTTAAGAAATGAGTTATCAGCTAGCGAGTTAGTCTGTCGA 622  
616 GGAGTCACTATCACTACATTTCCCAAAAGCGATTTATCTGCTAGTGAATTAGCAGCAGCT 675  
623 GAAGCTATTGGAATGGGAAGCAGGATCTCGTCTCTTCAAGTTCTAGTTATTAATGCA 682  
676 AAGCACATCTGCTGGTGGAAAAAATATGCAACCGAGTCAAGTTAAGCTATTCTTCAACAGCT 735  
683 AATCCAGCTCAACCAAGATTGTCAGAGAACCAATCTGACTGTCTCACTCCAACTTATCAT 742  
736 AGTGACAAATAACAGCAATCTGTAGCAAA-----AGGATCAACTAGCAGCCA 783  
743 CAAATCAAGGGGAAACAAATTTCAAGCCTTTTACGTAATTTGATGCTAAGACCTTATCA 802  
784 GCAATATAATCTGAAATCTCCAGAGTCTTTTGAAGGAATCTATGATTCACCTAGCGCC 843  
803 GAACGCCATGTGGAATCTGATGGCTTTATTTTCGACCCAGCGCAATCAAGTCCGAACC 862  
844 CAACGTTACAGTGAATCAGATGGCTGGTCTTTGACCTGCTGAAGATTATCAGTCTGTACA 903  
863 GCCAGAGGTGATGCTGCTCCATCATGTAACCATTAACCTTTATCCCTTATGAACAAATG 922  
904 CCAATGAGGTTGCGAATTCGCGATGGCGACCAATTAACCTTTATTCCTTACAGCAAGCTT 963  
923 TCTGAATTTGAAAAACGAATTTGCTGCTATTATTTCCTTCCTTCGT 964  
964 TCTGCTTAGAAGAAAGATTGCCAGATGGTGGCTTATCAGT 1005

RESULT 22  
US-09-769-744A-23  
; Sequence 23, Application US/09769744A  
; Publication No. US20030134407A1  
; GENERAL INFORMATION:  
; APPLICANT: Le Page, Richard WF  
; APPLICANT: Wells, Jeremy M  
; APPLICANT: Hanniffy, Sean B  
; APPLICANT: Hansbro, Philip M  
; TITLE OF INVENTION: Proteins  
; FILE REFERENCE: PWC/P21122WO  
; CURRENT APPLICATION NUMBER: US/09/769,744A  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: PCT/GB99/02452  
; PRIOR FILING DATE: 1999-07-27  
; PRIOR APPLICATION NUMBER: GB 9816336.3  
; PRIOR FILING DATE: 1998-07-27  
; PRIOR APPLICATION NUMBER: US 60/125329  
; PRIOR FILING DATE: 1999-03-19  
; NUMBER OF SEQ ID NOS: 196  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 1455  
; TYPE: DNA

RESULT 23

383 ATTCGGACAAAAGAGATTAAACGTCAGAGCAGGAAACGAGTCATAATCATTAATCA 442  
439 GTTCGAACATAAGATGAATCAATCGTCAAAACAAGAAC---ATGTCAAAGATATGAG 495  
443 AGACGAGATATGCTGTTGCTGAGCCAGAGCCCAAGACGTTATACACGAGTATGAG 502  
496 AAGGTTAACTCTAATGTTGCTGTAGCAAGGCTCAGGGACGATATACGCAAAATGATG 555  
503 TATATCTTCAATGATCTGATATCATTTGAGGACAGGGGATGCTTATATCGTTCTCTC 562  
556 TATGCTTTTAAATCCAGCTGATATATCGAAGATACGGGTAATGCTTATATCGTTCTCT 615  
563 GCGACCAATACCAATACATTCCTTAAGATGAGTTATCAGTCAGCGAGTATGCTGCTG 622  
616 GGAGTCACTATCACTACATTCCTCAAAAGCGATTTATCTGCTAGTGAATTAGCAGC 675  
623 GAAGCTTATTTGAATGGGAAGCAGGATCTCGTCTCTTTCAGTTTCAAGTTCTAGTT 682  
676 AAGGACATCTGCTGGGAAAAAATATGCAACGAGTCAGTTAAGCTATTCTTCAACAG 735  
683 AATCCAGCTCAACCAAGATTGTCAGAGAACCAATCTGACTCTCACTCCAACTTATCAT 742  
736 AGTGACATAACACGCAATCTGTAGCAA-----AGATCAACTAGCAAGCCA 783  
743 CAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGAATGATGCTAAACCTTATCA 802  
784 GCAATATAATCTGAAAAATCTCCAGAGTCCTTTGAAGGAACTCTATGATTCACCTAG 843  
803 GAAGCGCATGTGGAATCTGATGCGCTTATTTTCAGCCAGCGGCAATACAGTCGAACC 862  
844 CAACGTTACGTGAATCAGATGCGCTGCTTTTGACCTGCGACCATTTACACITTTAT 903  
863 GCAGAGGTGTAGTGTCTCTCATGTTAACAATACCACTTATCCCTTATGAACAATG 922  
904 CCAATGCGAGTTCGATTCGATGCGGACGATGCGCTGCTTTGACCTGTAAGATTAT 964  
923 TCTGAATGGAAGAAACGAATGCTGCTGATTTATTTCCCTTCCT 964  
964 TCTGCTTTAGAGAAAGAAATGCGAGAAATGCGTATGCTGCTATCACT 1005

RESULT 24  
US-10-412-850-7  
; Sequence 7, Application US/10412850  
; Publication No. US20040001836A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; TITLE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-686  
; CURRENT APPLICATION NUMBER: US/10/412,850  
; CURRENT FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 09/468,656  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 7  
; LENGTH: 1455  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-10-412-850-7

Query Match 17.5%; Score 401; DB 16; Length 1455;  
Best Local Similarity 65.7%; Pred. No. 1.1e-101;  
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

QY 23 CACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGTTTCTTATATAGATGGTATCAG 82  
DB 79 CAGCATGTTGCGAGGAAATAAGGACAAATATCGTCTCTTATGTTGATGGCAGCCAG 138  
QY 83 GCTGGTCAAAAGCAGAAACTTTCACACAGATCAAGTCAAGAGAGGGGATCAAC 142  
DB 139 TCAAGTCAGAAAGTGAAGAACTTTCACACAGATCAAGTCAAGAGAGGAAATTCAG 198  
QY 143 GCCGAAACAAATGTTNATCAAGATTACGATCAAGGTTATGTCCTCTTCATGGAGACCAT 202  
DB 199 GCTGAGCAAAATGTTAATCAAAATTCAGATCAGGCTATGTAACGTCACAGGTGACCCAC 258  
QY 203 TATCATTAATAATGGAAGGTTCTCTTATGATGCTATCATGAGAGTCTCTCATG 262  
DB 259 TATCATTAATAATGGAAGGTTCTCTTATGATGCTCTTATGAGAGAACTCTGTATG 318  
QY 263 AAAGATCGAATATCACTTGAAGATTCAGACATTTGCAATGAATCAAGGTTGTTAT 322  
DB 319 AAGGATCAAACTATCAACTTAAAGACGCTGATTTGCAATGAATCAAGGTTGTTAT 378  
QY 323 GTCATTAAGTAAAGGGTAAATATCTATGTTACCTTAAAGGATGTCAGCTCATCGGATAAT 382  
DB 379 ATCATCAAGGTCGAGGAAATATTTATGTTCTACTGAAAGATGTCAGCTCATGCTGATAAT 438

QY 383 ATTCGGACAAAAGAGATTAAACGTCAGAGCAGGAAACGAGTCATAATCATTAATCA 442  
DB 439 GTTCGAACATAAGATGAATCAATCGTCAAAACAAGAAC---ATGTCAAAGATATGAG 495  
QY 443 AGACGAGATATGCTGTTGCTGAGCCAGAGCCCAAGACGTTATACACGAGTATGAG 502  
DB 496 AAGGTTAACTCTAATGTTGCTGTAGCAAGGCTCAGGGACGATATACGCAAAATGATG 555  
QY 503 TATATCTTCAATGATCTGATATCATTTGAGGACAGGGGATGCTTATATCGTTCTCTC 562  
DB 556 TATGCTTTTAAATCCAGCTGATATATCGAAGATACGGGTAATGCTTATATCGTTCTCT 615  
QY 563 GCGACCAATACCAATACATTCCTTAAGATGAGTTATCAGTCAGCGAGTATGCTGCTG 622  
DB 616 GGAGTCACTATCACTACATTCCTCAAAAGCGATTTATCTGCTAGTGAATTAGCAGC 675  
QY 623 GAAGCTTATTTGAATGGGAAGCAGGATCTCGTCTCTTTCAGTTTCAAGTTCTAGTT 682  
DB 676 AAGGACATCTGCTGGGAAAAAATATGCAACGAGTCAGTTAAGCTATTCTTCAACAG 735  
QY 683 AATCCAGCTCAACCAAGATTGTCAGAGAACCAATCTGACTCTCACTCCAACTTATCAT 742  
DB 736 AGTGACATAACACGCAATCTGTAGCAA-----AGATCAACTAGCAAGCCA 783  
QY 743 CAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGAATGATGCTAAACCTTATCA 802  
DB 784 GCAATATAATCTGAAAAATCTCCAGAGTCCTTTGAAGGAACTCTATGATTCACCTAG 843  
QY 803 GAAGCGCATGTGGAATCTGATGCGCTTATTTTCAGCCAGCGGCAATACAGTCGAACC 862  
DB 844 CAACGTTACGTGAATCAGATGCGCTGCTTTTGACCTGTAAGATTATCAGTCGTACA 903  
QY 863 GCAGAGGTGTAGTGTCTCTCATGTTAACAATACCACTTATCCCTTATGAACAATG 922  
DB 904 CCAATGCGAGTTCGATTCGATGCGGACGATGCGCTGCTTTGACCTGTAAGATTAT 964  
QY 923 TCTGAATGGAAGAAACGAATGCTGCTGATTTATTTCCCTTCCT 964  
DB 964 TCTGCTTTAGAGAAAGAAATGCGAGAAATGCGTATGCTGCTATCACT 1005

RESULT 25  
US-10-387-783-7  
; Sequence 7, Application US/10387783  
; Publication No. US20040005331A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; TITLE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-683  
; CURRENT APPLICATION NUMBER: US/10/387,783  
; CURRENT FILING DATE: 2003-03-13  
; PRIOR APPLICATION NUMBER: 09/468,656  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 7  
; LENGTH: 1455  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-10-387-783-7

Query Match 17.5%; Score 401; DB 16; Length 1455;  
Best Local Similarity 65.7%; Pred. No. 1.1e-101;  
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

QY 23 CACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGTTTCTTATATAGATGGTATCAG 82  
DB 79 CAGCATGTTGCGAGGAAATAAGGACAAATATCGTCTCTTATGTTGATGGCAGCCAG 138  
QY 83 GCTGGTCAAAAGCAGAAACTTTCACACAGATCAAGTCAAGAGAGGGGATCAAC 142  
DB 139 TCAAGTCAGAAAGTGAAGAACTTTCACACAGATCAAGTCAAGAGAGGAAATTCAG 198  
QY 143 GCCGAAACAAATGTTNATCAAGATTACGATCAAGGTTATGTCCTCTTCATGGAGACCAT 202  
DB 199 GCTGAGCAAAATGTTAATCAAAATTCAGATCAGGCTATGTAACGTCACAGGTGACCCAC 258  
QY 203 TATCATTAATAATGGAAGGTTCTCTTATGATGCTATCATGAGAGTCTCTCATG 262  
DB 259 TATCATTAATAATGGAAGGTTCTCTTATGATGCTCTTATGAGAGAACTCTGTATG 318  
QY 263 AAAGATCGAATATCACTTGAAGATTCAGACATTTGCAATGAATCAAGGTTGTTAT 322  
DB 319 AAGGATCAAACTATCAACTTAAAGACGCTGATTTGCAATGAATCAAGGTTGTTAT 378  
QY 323 GTCATTAAGTAAAGGGTAAATATCTATGTTACCTTAAAGGATGTCAGCTCATCGGATAAT 382  
DB 379 ATCATCAAGGTCGAGGAAATATTTATGTTCTACTGAAAGATGTCAGCTCATGCTGATAAT 438

Query Match 17.5%; Score 401; DB 10; Length 3120;  
Best Local Similarity 65.7%; Pred. No. 1.9e-101;  
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

Y 83 GCTGGTCAAAAGGCGAGAACTTGCACACAGATGAAGTCAGTAAGAGGAGGGGATCAAC 142  
b 139 TCAAGTCAGAAAAGTGAATCTTGACACAGACAGGTTAGCCAGAAAGAGAAATTCAG 198  
Y 143 GCCGAACAAATNGTATCAAGATTACGATTCAGGTTATGTGACCTCTCATGAGACCAT 202  
b 199 GCTGAGCAATTTGTAATCAAAATTAACAGATCAGGCTATGTAACGTCAACGCTGACCA 258  
Y 203 TATCATTTACTATATGCGAAGGTTCTTATGATGCTATCATCATCAGTGAAGAGCTCTCATG 262  
b 259 TATCATTTACTATATGCGAAGGTTCTTATGATGCTCTTATGATGAGAACTCTTTGATG 318  
Y 323 GTCAATTAAGGTAACCGTAAATPACTATGTTTACCTTAAGGATGCGAGCTCATGCGGATAAT 382  
b 379 ATCATCAGGTCGATGGAATATTTATGTTCTACCTGAAAGATGCGAGCTCATGCTGATAAT 438  
Y 383 ATTCGGAACAAAGAGAGATTAAAGTCAAGAGCAGGACGAGTCATTAATCATTAATCTCA 442  
b 439 GTTCGAACTTAAGATGAATCAATCGTCAAAAACRAAGAAC---ATGTCAAAGATATGAG 495  
Y 443 AGACGAGATATGCTGTTGCTGCGACCGACAGCCAGGAGCTTATACAAACGAGATGATGG 502  
b 496 AAGGTTAACTCTAAATGTTGCTGTAGCAAGGCTCTCAGGAGCATATACGACAAATGATGT 555  
Y 503 TATATCTTCAATGCACTCTGATATCAATGAGGACACCGGTCGATCTTATATCGTTCTCTAC 562  
b 556 TATGTTCTTATCCAGCTGATATATCGAGATACCGGTAATGCTTATATCGTTCTCTCAT 615  
Y 563 GCGGACCAATTACCATTTACTTAAGATGATTAATGATGATGATGATGATGATGATGATG 622  
b 616 GGAGGTCACATCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 675  
Y 623 GAAGCCTATTGGAATGGAGGAGGAGGATCTCGTCTCTTCTCAAGTCTTAGTTTATATGCA 682  
b 676 AAAGCACATCTGGCTGGAAAAAATATGCAACCGAGTCAGTTAGCTATTCTTCAACAGCT 735  
Y 683 AATCCAGCTCAACCAAGATTTGTGAGAGAACCAATCTGACTGTCTCACTCCAACTTATCAT 742  
b 736 AGTGACAATAACACGCAATCTGTAGCAAA-----AGGATCAACTAGCAAGCCA 783  
Y 743 CAAAATCAAGGGGAAAAACATTTCAAGCCCTTTTACGTAATGTTATGCTAAACCCCTTATCA 802  
b 784 GCAATTAATCTGAAAAATCTCCAGAGTCTTTTGAAGGAACTCTTATGATTCACCTAGCGCC 843  
Y 803 GAACGCCATGTGGAATCTGATGCGCTTTATTTTCGACCGCAGGCAATCAACAGTCGAAAC 862  
b 844 CAACGTTACAGTGAATCAGATGCGCTGGTCTTTGACCCCTGCTAAGATTATCAGTCGTACA 903

Y 83 GCTGGTCAAAAGGCGAGAACTTGCACACAGATGAAGTCAGTAAGAGGAGGGGATCAAC 142  
b 139 TCAAGTCAGAAAAGTGAATCTTGACACAGACAGGTTAGCCAGAAAGAGAAATTCAG 198  
Y 143 GCCGAACAAATNGTATCAAGATTACGATTCAGGTTATGTGACCTCTCATGAGACCAT 202  
b 199 GCTGAGCAATTTGTAATCAAAATTAACAGATCAGGCTATGTAACGTCAACGCTGACCA 258  
Y 203 TATCATTTACTATATGCGAAGGTTCTTATGATGCTATCATCATCAGTGAAGAGCTCTCATG 262  
b 259 TATCATTTACTATATGCGAAGGTTCTTATGATGCTCTTATGATGAGAACTCTTTGATG 318  
Y 323 GTCAATTAAGGTAACCGTAAATPACTATGTTTACCTTAAGGATGCGAGCTCATGCGGATAAT 382  
b 379 ATCATCAGGTCGATGGAATATTTATGTTCTACCTGAAAGATGCGAGCTCATGCTGATAAT 438  
Y 383 ATTCGGAACAAAGAGAGATTAAAGTCAAGAGCAGGACGAGTCATTAATCATTAATCTCA 442  
b 439 GTTCGAACTTAAGATGAATCAATCGTCAAAAACRAAGAAC---ATGTCAAAGATATGAG 495  
Y 443 AGACGAGATATGCTGTTGCTGCGACCGACAGCCAGGAGCTTATACAAACGAGATGATGG 502  
b 496 AAGGTTAACTCTAAATGTTGCTGTAGCAAGGCTCTCAGGAGCATATACGACAAATGATGT 555  
Y 503 TATATCTTCAATGCACTCTGATATCAATGAGGACACCGGTCGATCTTATATCGTTCTCTAC 562  
b 556 TATGTTCTTATCCAGCTGATATATCGAGATACCGGTAATGCTTATATCGTTCTCTCAT 615  
Y 563 GCGGACCAATTACCATTTACTTAAGATGATTAATGATGATGATGATGATGATGATGATG 622  
b 616 GGAGGTCACATCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 675  
Y 623 GAAGCCTATTGGAATGGAGGAGGATCTCGTCTCTTCTCAAGTCTTAGTTTATATGCA 682  
b 676 AAAGCACATCTGGCTGGAAAAAATATGCAACCGAGTCAGTTAGCTATTCTTCAACAGCT 735  
Y 683 AATCCAGCTCAACCAAGATTTGTGAGAGAACCAATCTGACTGTCTCACTCCAACTTATCAT 742  
b 736 AGTGACAATAACACGCAATCTGTAGCAAA-----AGGATCAACTAGCAAGCCA 783  
Y 743 CAAAATCAAGGGGAAAAACATTTCAAGCCCTTTTACGTAATGTTATGCTAAACCCCTTATCA 802  
b 784 GCAATTAATCTGAAAAATCTCCAGAGTCTTTTGAAGGAACTCTTATGATTCACCTAGCGCC 843  
Y 803 GAACGCCATGTGGAATCTGATGCGCTTATTTTCGACCGCAGGCAATCAACAGTCGAAAC 862  
b 844 CAACGTTACAGTGAATCAGATGCGCTGGTCTTTGACCCCTGCTAAGATTATCAGTCGTACA 903  
Y 863 GCGAGAGGTGATGCTGCTCCCTCATGTTGATCAATTAACCATTTATCCCTTATGAACAAATG 922  
b 904 CCAATATGAGTTGCGAATCCGATGCGAGCAATTAACCATTTATTCCTTACAGCAGCTT 963  
Y 923 TCTGAATTTGAAAAACGAATTTGCTGCTTATTTATTCCTCTCGT 964  
b 964 TCTGCTTAGAAGAAAGATTGCGAAGTGGTGCCTATCAGT 1005

RESULT 26  
US-09-884-465A-1  
; Sequence 1, Application US/09884465A  
; Publication No. US2003007293A1  
; GENERAL INFORMATION:  
; APPLICANT: Shire Biochem, Inc.  
; APPLICANT: Hamel, Josee  
; APPLICANT: Brodeur, Bernard  
; APPLICANT: Martin, Denis  
; APPLICANT: Charland, Nathalie  
; APPLICANT: Ouellet, Catherine

QY	863	GC	CAGAGG	TG	TAG	TG	CC	CT	CG	TG	AA	CA	AT	TAC	CA	CT	TAT	CC	CT	TAT	GC	AA	CA	AT	G	922
DB	904	CC	AA	TG	GA	TT	GG	GA	TT	TC	GC	AT	GG	GA	CA	AT	TAC	CA	CT	TAT	TCC	TAT	TCC	TAT	GAC	963
QY	923	TC	TG	AA	TT	GG	AA	AA	CA	GA	AT	TG	CT	GA	TT	AT	TCC	CT	TC	GT					964	
DB	964	TC	TG	CT	TT	TAG	AA	AA	AA	AG	AT	TG	CA	GA	AT	GG	TG	CC	TA	TC	AG				1005	

```

RESULT 27
US-09-884-465A-2
/ Sequence 2, Application US/09884465A
/ Publication No. US20030077293A1
/ GENERAL INFORMATION:
/ APPLICANT: Shire Biochem, Inc.
/ APPLICANT: Hamel, Josee
/ APPLICANT: Brodeur, Bernard
/ APPLICANT: Martin, Denis
/ APPLICANT: Charland, Nathalie
/ APPLICANT: Ouellet, Catherine
/ TITLE OF INVENTION: Streptococcus Antigens
/ FILE REFERENCE: 055190-0044
/ CURRENT APPLICATION NUMBER: US/09/884,465A
/ CURRENT FILING DATE: 2001-06-20
/ PRIOR APPLICATION NUMBER: 60/212,683
/ PRIOR FILING DATE: 2000-06-20
/ NUMBER OF SEQ ID NOS: 384
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 5048
/ TYPE: DNA
/ ORGANSIM: Streptococcus pneumoniae
/ US-09-884-465A-2

```

Query Match 17.5%; Score 401; DB 10; Length 5048;  
Best Local Similarity 65.7%; Pred. No. 2.6e-101;  
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

Qy	23	CACCAAGCTGGT	CAGGTTAAGAAAGAGCTCTAAATCGAGTTTCTTATATAGATGGTGATCAG	82
Db	1855	CAGCATCGTTCCG	CAGGAAAATAAGGACAAATAATCGTCTCTTATGTGGATGGCAGCGCAG	1914
Qy	83	GCTGGTCAA	AAGGCAGAAAACTTCACACCCAGATCAAGTCAGTAAGAGGAGGGGATCAAC	142
Db	1915	TCAAGTCAGAAA	AAAGTGAACCTTCACACCCAGCCAGGTAGCCAGAAAGAAATTCAG	1974
Qy	143	GCCGAACAAA	TNGTNATCAAGATTACGGATCAAGGTTATGTGACCTCTCATGGAGACCAT	202
Db	1975	GCTGAGCAA	ATTGTAATCAAAATTTACAGATCAGGCGTATGTAAAGCTCACACGGTGACCC	2034
Qy	203	TATCATTTACT	ATAATGGCAAGGTTCCCTTATGATGCCATCATCAGTCAGGAAGCTCCTCATG	262
Db	2035	TATCATTTACT	ATAATGGGAAAGTTCCTTATGATGCCCTCTTTAGTGAAGAACTCTTGATG	2094
Qy	263	AAAGATCCGA	ATTATCAGTTCGAAGGATTACAGACATTTGTCAATGAAATCAAGGGTGCTTAT	322
Db	2095	AAGGATCCAA	ACTATCAACTTTAAGACCGCTGATTTGTCAATGAAGTCAGGGTGCTTAT	2154
Qy	323	GTCAATTAAG	GGTAAACGGTAAATATCTATGTTNATACCTTAAAGGATGCAGCTCATCGCGATTAAT	382
Db	2155	ATCATCAAGG	TCGATGGAAAATATTTATGCTTACCTGAAAGATGCAGCTCATGCTGATAAT	2214
Qy	383	ATTCCGCAAAA	AGAAAGAGATTAAACGTCAGAGCAGGAAACCGCAGTCTAATCATTAACCTCA	442
Db	2215	GTTCGAAC	TAAAGATGAAATCAATCGTCAAAAACCAAGAAC---ATGTCAAAAGATAATGAG	2271
Qy	443	AGAGCAGAT	ATAATGCTGTGTCGACGCCAGAGCCCAAGGACGTTATACAAACGATGATGGG	502
Db	2272	AAGGTTAA	CTCTAAATGTTGCTGTAGCAAGGCTCTAGGGAAGATATACGACAAATGATGGT	2331
Qy	503	TATATCTTT	CAATGCATCTGATATCAATTGAGGACACGGGTGATGCTTATATCGTTCCTCAAC	562

Db	2332	TATGTCCTTTAAATCCAGCTGATATATTCGAAGATACGGGTAATGCTTATATATCGTTCCTCAT	2339
Qy	563	GGGACCAATTACCATTACATTTCTTAAAGATGAGTTATCAGCTACGAGGTAGCTGTGCGCA	622
Db	2392	GGAGGTCACTATCACTACATTTCCCAAAGCGATTTATCTGCTAGTGAATTAGCAGCAGCT	2451
Qy	623	GAAGCCCTATTGGAATGGGAAGCAGGGAATCTCGTCTCTTTCAAAGTTCCTAGTTATAATGCA	682
Db	2452	AAAGCACATCTGGCTGGAAAAAATATGCAACCGAGTCAGTTAGCTATTTCTTCAACAGCT	2511
Qy	683	AATCCAGCTCAACCAAGATTGTCCAGAGAACCACAATCTGACTGTCACTCCAACTTATTCAT	742
Db	2512	AGTCACAAATACACGCAATCTGTAGCAA-----AGGATCAACTAGCAAGCCA	2559
Qy	743	CAAAATCAAGGGGAAAAATTTCAAGCCTTTTACGTGAATGTATGTCTAAACCCCTTATCA	802
Db	2560	GCAAAATAATCTGAAATCTCCAGAGCTTTTGAAGGAATCTATGATTTACCTTAGCGCC	2619
Qy	803	GAACGCCATGTGGAATCTGATGGCTTATTTTCGACCCAGCGGGAATCAACAAGTCGAACC	862
Db	2620	CAACGTTACAGTGAATCAGATGGCGCTGTCTTTGACCCCTGCTAAGAGTTATCAGTCGTACA	2679
Qy	863	GCCAGAGGTGATGTGTCCTTCATCGGTAAACAATTACCACTTTATCCCTTATGACAAATG	922
Db	2680	CCAAATGGAGTTGCGATTCGTCATGGGACCAATTACCACCTTAATTCCTTACGCAAGCTT	2739
Qy	923	TCTGAAATTTGAAAAACGAATTTGCTCGTATTAATTTCCCTTCGT	964
Db	2740	TCTGCTTTTAGAAGAAAAAGATTGGCAGAATGGTGCTATCAGT	2781

RESULT 28

US-10-158-844-192/c  
; Sequence 192, Application US/10158844

GENERAL INFORMATION:

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key

CITY: Rockville

STATE: Maryland

COUNTRY: U

ZIP: 20850

COMPUTER READABLE FORM

; MEDIUM TYPE: CD-R

```
COMPUTER: Dell Latitude Pent: 2.8
```

```

;
OPERATING SYSTEM: Windows 98

```

```

;
SOFTWARE: ASCII Text
*****

```

; CURRENT APPLICATION DATA: US/10/158 84  
; APPLICATION NUMBER: US/10/158 84

APPLICATION NUMBER: US/10/158,841  
 FILING DATE: 03-10-2002

FILING DATE: 03-JUN-2002  
CLASSIFICATION: UNKOWN

CLASSIFICATION: {UNKNOWN}  
PRIOR APPLICATION DATA:

; PRIOR APPLICATION DATA:   
 ; APPLICATION NUMBER: IIS 08/961-52

APPLICANT NUMBER: 03 00/2017,32  
FILING DATE: 1997-10-30

FILED DATE: 10/27/2000  
APPLICATION NUMBER: US 60/029,966

AFFIDAVIT NUMBER: 00 000 025130  
 FILING DATE: 1996-10-31

FILED DATE: 1990 10 31  
ATTORNEY/AGENT INFORMATION:

NAME: Hyman, Mark J.

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB340P1

INFORMATION FOR SEQ ID NO: 192:

SEQUENCE CHARACTERISTICS:

LENGTH: 6867 base pairs

TYPE: nucleic acid

```

;
; STRANDEDNESS: double

```

TOPOLOGY: linear

```

;
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 192:

```

US-10-158-844-192

Query Match 17.5%; Score 401; DB 13; Length 6867;  
Best Local Similarity 65.7%; Pred. No. 3.2e-101;  
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;  
23 CACCAAGCTGGTACAGTTAAGAAAGAGTCTAATCGAGTTCTTATAGATGGTATGATCAG 82  
6707 CAGCATCGTTCGAGGAAATTAAGCAATTAATCGTGTCTCTTATGTGATGCGCCAG 6648  
83 GCTGTGTCACAAAGGCGAAGAACTTTGACACAGATGAAGTCAAGTAAGAGGAGGAGGATCAAC 142  
6647 TCAAGTCAGAAAGTGAAACCTTGACACAGACAGCAGGTTAGCCAGAAAGAGAAATTCAG 6588  
143 GCGCAACAAATGATNATCAAGATTACGATCAAGGTATATGTGACCTCTCATGAGACCAT 202  
6587 GCTGAGCAATTTGTAATCAAAATACAGATCAGGGCTATGTAAAGTCACACGGTGACCA 6528  
203 TATCAATTAATTAAGGCAAGGTTCTTATGATGATGATGATGATGATGATGATGATGATG 262  
6527 TATCAATTAATTAAGGCAAGGTTCTTATGATGATGATGATGATGATGATGATGATG 6468  
263 AAGATCCGAATATCAGTTGAGGATTCAGACATGTCATGATGATGATGATGATGATGATG 322  
6467 AAGGATCCAAATCATCACTTAAAGACGCTGATGATGATGATGATGATGATGATGATG 6408  
323 GTCAATTAAGGTAACCGTAAATTAATGATGATGATGATGATGATGATGATGATGATG 382  
6407 ATCAATCAAGTTCGATGAAATTAATGATGATGATGATGATGATGATGATGATGATG 6348  
383 ATTCGACAAAGAGAGATTAAGTCAAGACGAGGAAAGCGAGTCAATATCAATCAATCA 442  
6347 GTTCGAATCAAGATGAATCAATCGTCAAAACAAAGAAC---ATGTCAAAGATTAATGAG 6291  
443 AGACGAGATTAATCTCTGTCGACGAGCCAGGAGGATGATGATGATGATGATGATGATG 502  
6290 AAGGTTAACTCTTAATGTTGTCGACAGGATCTCAGGAGGATGATGATGATGATGATG 6231  
503 TATATCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 562  
6230 TATGCTTTAATCCAGCTGATATTCGAGATGATGATGATGATGATGATGATGATGATG 6171  
563 GCGCAGCAATTAATCAATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 622  
6170 GGAGGTCATCTACATCAATTCGCAAAAGGATTTATCTGCTAGTGAATTAGCAGCAGCT 6111  
623 GAGCCTATGTAATGGAGAGGAGGATCTGCTCTCTTCAAGTCTGATGATGATGATGATG 682  
6110 AAGACCATCTGCTGGGAAATTAATGCAACCGAGTCAGTAAAGCTATCTTCAACAGCT 6051  
683 AATCCAGCTCAACCAAGATGTCAGAGAACCAATCTGATGATGATGATGATGATGATGAT 742  
6050 AGTGACATTAACAGGATCTGTAGCAAA-----AGGATCAACTAGCAGCA 6003  
743 CAAATCAAGGGGAAACATTTCAAGCCTTTTACGTAATGATGATGATGATGATGATGATG 802  
6002 GCAATTAATCTGAAATCTCCAGAGCTTTTGAAGGAACTCTATGATTTCACTAGCGCC 5943  
803 GAAGCCATGTAATCTGATGCTTATTTTCAAGCCAGGCAATCAAGATGCAAGC 862  
5942 CAACGTTACAGTGAATCAGATGCTGCTCTTTGACCCCTGCTAAGATTAATCAGTCGTAC 5883  
863 GCCAGAGGTCAGTGTCCCTCATGGTAAACCATTAACCACTTTATCCCTTATGAACAAATG 922  
5882 CCAATGGAGTTGGATTCGCAATGGGAGGATTCACCACTTTATCCCTTACAGCAAGCTT 5823  
923 TCTGAATGGAAAAACGAATGCTGATTAATTTCCCTTCGT 964  
5822 TCTGCTTAGAAGAAAGATTCGCAATGCTGCTATCAGT 5781

RESULT 29

US-09-884-465A-257

; Sequence 257, Application US/09884465A

; Publication No. US2003007293A1

GENERAL INFORMATION:  
; APPLICANT: Shire Biochem, Inc.  
; APPLICANT: Hamel, Josee  
; APPLICANT: Brodeur, Bernard  
; APPLICANT: Martin, Denis  
; APPLICANT: Charland, Nathalie  
; APPLICANT: Ouellet, Catherine  
; TITLE OF INVENTION: Streptococcus Antigens  
; FILE REFERENCE: 055190-0044  
; CURRENT APPLICATION NUMBER: US/09/884,465A  
; CURRENT FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: 60/212,683  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 384  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 257  
; LENGTH: 819  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Unknown Organism  
US-09-884-465A-257

Query Match 16.2%; Score 370; DB 10; Length 819;  
Best Local Similarity 73.2%; Pred. No. 4.2e-93;  
Matches 534; Conservative 0; Mismatches 130; Indels 66; Gaps 2;  
1492 GCAATTAACCTACACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1551  
Db 3 GCAATTAACCTACACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 62  
1552 AGAAGACGGTTATATCTTTGATCCTCGTGATATACCAAGTGTAGGGGGATGCTTATGT 1611  
Db 63 AGAAGACGGTTATATCTTTGATGATGATGATGATGATGATGATGATGATGATGATG 90  
1612 AACTCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1671  
Db 91 TGGATTAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 125  
1672 AGCGCAGCCAGGCTTATGCTTAAGAGAAAGGTTTGACCCCTCCCTTCGACAGACCATCA 1731  
Db 126 AGCGCAGCCAGGCTTATGCTTAAGAGAAAGGTTTGACCCCTCCCTTCGACAGACCATCA 185  
1732 GGATTCAGAAATCTGAGGCAAAAGGAGCAGAAAGCTATCTCAACCCGCTGAAAGCAGC 1791  
Db 186 GGATTCAGAAATCTGAGGCAAAAGGAGCAGAAAGCTATCTCAACCCGCTGAAAGCAGC 245  
1792 TAAGAAAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1851  
Db 246 TAAGAAAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 305  
1852 CGGTAGTTTATCATACCTCATTATGACCATTAACCATTAACCATTAACCATTAACCATTA 1911  
Db 306 CGGTAGTTTATCATACCTCATTATGACCATTAACCATTAACCATTAACCATTAACCATTA 365  
1912 CGAAGCCCTTTATGAGGCACTTAAGGGGTATAGTCTTTGAGGATCTTTTGGCGAGTGTCAA 1971  
Db 366 CGAAGCCCTTTATGAGGCACTTAAGGGGTATAGTCTTTGAGGATCTTTTGGCGAGTGTCAA 425  
1972 GTACTATGTCGAACATCCAAACGAAAGCTCGCATTAATGATGATGATGATGATGATGATG 2031  
Db 426 GTACTATGTCGAACATCCGCGAAAGCTAGT-----GACCATGTTTCTGTAAGAAATAGGC 476  
2032 CGACCATGTTCAAGAAACAAANAATGTCAGCTGATGATGATGATGATGATGATGATGATG 2091  
Db 477 AGACCAAGATAGTAAACCTGATGAAGATGAAGAAACATGATGAAGATGATGATGATGATG 536  
2092 CGAGGAGAAACCTCAGACAGAAACAAACCTGAGGAAAGAAACCCCTCGAGAGAGAAACCGCA 2151  
Db 537 CCTGAATCTGATGAAGAAAGAGATCACGCTGTTTAAATCCTTCAGCAGATATATCTTA 596  
2152 AAGCGAGAAACCAAGAGTCTTCCAAACCAACAGAGGAAACCAAGAGAAATCACCGAGGAATC 2211

Db	597	TAACCAAGCACTGATACGGAAGACACAGAGAAAGTGAAGATACCAAGATGAGGC	656
Qy	2212	AGAAGAACCT	2221
Db	657	TGAAATTCCT	666

```

RESULT 30
US-10-324-143-10
/ Sequence 10, Application US/10324143
/ Publication No. US20030232976A1
/ GENERAL INFORMATION:
/ APPLICANT: HAMEL, JOSEF
/ APPLICANT: CHARLAND, NATHALIE
/ APPLICANT: BRODEUR, BERNARD R.
/ APPLICANT: MARTIN, DENIS
/ APPLICANT: BLAIS, NORMAND
/ APPLICANT: OVELETTE, CATHERINE
/ TITLE OF INVENTION: STREPTOCOCCUS ANTIGENS
/ FILE REFERENCE: 55190-56
/ CURRENT APPLICATION NUMBER: US/10/324,143
/ CURRENT FILING DATE: 2002-12-20
/ PRIOR APPLICATION NUMBER: 60/341,252
/ PRIOR FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 160
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 10
/ LENGTH: 819
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artific
/ OTHER INFORMATION: modified Streptococcus
US-10-324-143-10

```

Query Match	16.2%	Score 370;	DB 16;	Length 819;
Best Local Similarity	73.2%	Pred. No. 4.2e-93;		
Matches 534;	Conservative	0;	Mismatches 130;	Indels 66;
				Gaps 2;

QY	1492	GCAAAATTACCTACACTGATGNTGAGATTCAAGTAGCCAAAGTTGGCGGCGAAGTACACAAC	1551
Db	3	GCAAAATTACCTACACTGATGNTGAGATTCAAGTAGCCAAAGTTGGCGGCGAAGTACACAAC	62
QY	1552	AGAAGACGGTTATATCTTTTGATCCTCGTGATATACCAAGTATGAGGGGGATGCCTATGT	1611
Db	63	AGAGACGGTTATATCTTTTGATCCTCGTGATATACCAAGTATGAGGGGGATGCCTATGT	90

QY 1612 ARCTCCACATATGACCCCATAGCCACTGGATTAAAAAAGATAGTTTGTCTGAAGCTGAGAG 1671

nb 91 TGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAG 125

QY 1672 AGCGCAGCCCGAGCTTATGCTTAAGAGAGAAAGTTTGACCCCTCCTTGCACAGACCATCA 1731

ph 126 AGCGCAGCCCGAGCTTATGCTTAAGAGAGAAAGTTTGACCCCTCCTTGCACAGACCATCA 185

QY	1732	GGATT CAGGAATACTGAGCAAAAGGAGCAGAAGCTATCTTCAAACCGCGTGAAGCAGC	1791
ph	186	GGATT CAGGAATACTGAGCAAAAGGAGCAGAAGCTATCTTCAAACCGCGTGAAGCAGC	245

QY	1792	TAGAAAGGTGCCACTTGATCGTATGCTTCAATCTTCAATATATCTGTGAAAGTCAAAA	1851
Dh	245	TATGTAAGGTGCCACTTGATCGTATGCTTCAATCTTCAATATATCTGTGAAAGTCAAAA	305

CGGTAGTTTAAATCATACCTCATTTATGACCAATTACCAATAACATCAAAATTGAGTGCTTTGA 1911

QY	1912	CGAAGGCCTTTATGAGGCACCTAAGGGGTATACTCTTGAGGATCMTTTGGCGACTGTCAA	1971
bt	366	CGMTCGGCCCTTTATGCGCAACCTACCGGGGCTATGCTTTGAGGATCMTTTGGCGACTGTCAA	425

QY 1972 GTACTATGTCGAACATCCAAACGAACGTCGCAATTGCTTTTGGTAACGCTAG 20311

Db	426	GTACTATGTCGAACCGGNA	CGCTAGT-----GACCATGTTCTGTAATAAAGGC	476
Qy	2032	CGACCAATGTTCAAAGAAA	CAAAAATCGTCAAGCTGATACCAATCAACGGAAAAACCAAG	2091
Db	477	AGACCAAGATAGTAAACCT	GATCAAGANTAGGAACATGATGAATAGTGAAGCAACTCA	536
Qy	2092	CGAGGGAACCTCAGACAGAAA	AAACCTGAGGAGAAACCCCTCGAGAGAGAGAAACCGCA	2151
Db	537	CCCTGGAATCTGATGAAA	AAAGAGAAATCACGCTGTTTAAATCCTTCAGCAGTAAATCTTTA	596
Qy	2152	AAGCGAACCAGAGTCTC	CAAAAACCAACAGAGGAACAGAGAATCACCAGGAATC	2211
Db	597	TAAACCAAGACTGTATAC	GGAAGAGACAGAGGAAGAGCTGAAGATACCACAGATGAGGC	656
Qy	2212	AGAAGAACCT	2221	
Db	657	TGAATTCCT	656	

```

RESULT 31
US-10-324-143-4
; Sequence 4, Application US/10324143
; Publication No. US20030232976A1
; GENERAL INFORMATION:
; APPLICANT: HAMEL, JOSEE
; APPLICANT: CHARLAUD, NATHALIE
; APPLICANT: BRODEUR, BERNARD R.
; APPLICANT: MARTIN, DENIS
; APPLICANT: BLAIS, NORMAND
; APPLICANT: OUELETTE, CATHERINE
; TITLE OF INVENTION: STREPTOCOCCUS ANTIGENS
; FILE REFERENCE: 55190-56
; CURRENT APPLICATION NUMBER: US/10/324,143
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: 60/341,252
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial
; OTHER INFORMATION: modified Streptococcus
; OTHER INFORMATION: sequence
US-10-324-143-4

```

	Query Match	16.1%;	Score 368.4;	DB 16;	Length 816;
	Best Local Similarity	73.0%;	Pred. No. 1.2e-92;		
	Matches 533;	Conservative 0;	Mismatches 131;	Indels 66;	Gaps 2;
QY	1492	GCAAAATTACCTTACACTGATGATGAGATTCAAGTAGTCCAAAGTTGCGACGCGAAGTACACAAAC	1551		
Db	3	GCAAAATTACCTTACACTGATGATGAGATTCAAGTAGTCCAAAGTTGCGACGCGAAGTACACAAAC	62		
QY	1552	AGAAGACGGTTATATCTTTTGATCCTCGTGATATATACCAAGTGTATGACGGGGATGCGTATGT	1611		
Db	53	ACACACACGGCTTATCTCTTTGATATATGTT	90		

QY 1492 GC A A T T A C C T A C T A C T G A T G A T C A G A T T C A A G T A G C C A A G T T G C A G G C A A G T A C A C A A C 1551

[illegible]

QY	1612	AACTCCACATATGACCCCATAGCCACTGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAG	1671
QY			
QY	01	TTGATTTAAAAAAGATAGTTTGTCTGAAGCTGAGAG	125

1672 AGCGCAGCCCGAGGCTTATGCTAAAGAGAAAGGTTTGACCCCTCTCTTCGACAGACCATCA 17311

QY	1732	GGATTTCAGGAATTA	CTGAGGCAAAAGG	CGACGAGCTT	CTCAACCGCGTGA	AAAGCAGC	1791
ST	180	CGGTTTCGCGTAT	CTCTCTGCTGCT	TAACCGACGAGCT	TCTTCAACCGCGTGA	AAAGCAGC	245

1792 TAAGAAAGGTGCCACTTGATCGTATGCCTTACAATCTTCAATATACTGTAGAGTCAAAA 1851







TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical  
TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro  
FILE REFERENCE: 12287.31  
CURRENT APPLICATION NUMBER: US/09/452,599  
CURRENT FILING DATE: 1999-12-01  
PRIOR APPLICATION NUMBER: 08/526,840  
PRIOR FILING DATE: 1995-09-11  
PRIOR APPLICATION NUMBER: 08/304,732  
PRIOR FILING DATE: 1994-09-12  
NUMBER OF SEQ ID NOS: 177  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 34

LENGTH: 841

TYPE: DNA

ORGANISM: Streptococcus pneumoniae

US-09-452-599-34

Query Match 11.9%; Score 272.4; DB 9; Length 841;  
Best Local Similarity 92.9%; Pred. No. 1.3e-65;  
Matches 325; Conservative 0; Mismatches 6; Indels 19; Gaps 3;  
QY 1956 TTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGAACTCCGCGATTCAGATAATG 2015  
DB 841 TTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGAACTCCGCGATTCAGATAATG 783  
QY 2016 GTTTTGGTAACCTAGCGACCATGTTTCAAGAAACAAAATGGTCAAGCTGATACCAATC 2075  
DB 782 GTTTTGGTAACCTAGCGACCATGTTTCAAGAAACAAAATGGTCAAGCTGATACCAATC 723  
QY 2076 AAACGGAAAAACCAACGAGGAGAACTCCAGAGAAACCTCAGACAGAAACCTGAGGAGAAACCCCTC 2135  
DB 722 AAACGGAAAAACCAACGAGGAGAACTCCAGAGAAACCTCAGACAGAAACCTGAGGAGAAACCCCTC 663  
QY 2136 GAGAAGAGAAACCGCAAGGAGAGAACTCCAGAGAAACCTCAGAGAAACCTCAGAGAAACCTCAGAGAA 2187  
DB 662 GAGAAGAGAAACCGCAAGGAGAGAACTCCAGAGAAACCTCAGAGAAACCTCAGAGAAACCTCAGAGAA 603  
QY 2188 -----ACCAGAGAACTCAGAGAGGCTGAAGATTTTCTTGGAAAAATCCAGGAT 2290  
DB 542 T---AAGAAACTCAGAGAGGCTGAAGATTTTCTTGGAAAAATCCAGGAT 496

RESULT 36

US-10-121-120-34/c  
; Sequence 34, Application US/10121120  
; Publication No. US20030180733A1  
; GENERAL INFORMATION:  
; APPLICANT: Bergeron, Michel G.  
; APPLICANT: Ouellette, Marc  
; APPLICANT: Roy, Paul H.  
; TITLE OF INVENTION: Specific and Universal Probes and Amplification  
; TITLE OF INVENTION: Primers  
; TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial  
; TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical  
; TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro  
; FILE REFERENCE: 12287.31  
; CURRENT APPLICATION NUMBER: US/10/121,120  
; CURRENT FILING DATE: 2002-04-11  
; PRIOR APPLICATION NUMBER: 09/452,599  
; PRIOR FILING DATE: 1999-12-01  
; PRIOR APPLICATION NUMBER: 08/304,732  
; PRIOR FILING DATE: 1994-09-12  
; NUMBER OF SEQ ID NOS: 177  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 841  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-10-121-120-34

Query Match 11.9%; Score 272.4; DB 15; Length 841;  
Best Local Similarity 92.9%; Pred. No. 1.3e-65;  
Matches 325; Conservative 0; Mismatches 6; Indels 19; Gaps 3;  
QY 1956 TTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGAACTCCGCGATTCAGATAATG 2015  
DB 841 TTTTGGCGACTGTCAGTACTATGTCGAACATCCAAACGAACTCCGCGATTCAGATAATG 783  
QY 2016 GTTTTGGTAACCTAGCGACCATGTTTCAAGAAACAAAATGGTCAAGCTGATACCAATC 2075  
DB 782 GTTTTGGTAACCTAGCGACCATGTTTCAAGAAACAAAATGGTCAAGCTGATACCAATC 723  
QY 2076 AAACGGAAAAACCAACGAGGAGAACTCCAGAGAAACCTCAGACAGAAACCTGAGGAGAAACCCCTC 2135  
DB 722 AAACGGAAAAACCAACGAGGAGAACTCCAGAGAAACCTCAGACAGAAACCTGAGGAGAAACCCCTC 663  
QY 2136 GAGAAGAGAAACCGCAAGGAGAGAACTCCAGAGAAACCTCAGAGAAACCTCAGAGAAACCTCAGAGAA 2187  
DB 662 GAGAAGAGAAACCGCAAGGAGAGAACTCCAGAGAAACCTCAGAGAAACCTCAGAGAAACCTCAGAGAA 603  
QY 2188 -----ACCAGAGAACTCAGAGAGGCTGAAGATTTTCTTGGAAAAATCCAGGAT 2290  
DB 542 T---AAGAAACTCAGAGAGGCTGAAGATTTTCTTGGAAAAATCCAGGAT 496

RESULT 37

US-09-769-736-23  
; Sequence 23, Application US/09769736  
; Publication No. US20030138775A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbial Technics Limited  
; APPLICANT: Le Page, Richard WF  
; APPLICANT: Wells, Jeremy M  
; APPLICANT: Hanniffy, Sean B  
; TITLE OF INVENTION: Proteins  
; FILE REFERENCE: PWC/P21089WO  
; CURRENT APPLICATION NUMBER: US/09/769,736  
; CURRENT FILING DATE: 2003-02-14  
; PRIOR APPLICATION NUMBER: GB 9816335.5  
; PRIOR FILING DATE: 1998-07-27  
; PRIOR APPLICATION NUMBER: US 60/125163  
; PRIOR FILING DATE: 1999-03-19  
; NUMBER OF SEQ ID NOS: 212  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 1146  
; TYPE: DNA  
; ORGANISM: Streptococcus agalactiae  
US-09-769-736-23

Query Match 9.8%; Score 225; DB 10; Length 1146;  
Best Local Similarity 63.5%; Pred. No. 3.7e-52;  
Matches 376; Conservative 0; Mismatches 183; Indels 33; Gaps 1;  
QY 100 AAACCTTGACACCATGATGAAGTCAGTAAGAGGAGGGGATCAACGCCGGAACAATNGTNA 159  
DB 162 AAACAAACAGATGGATCAATCAAGTGTCTGAAGAGGATCTCTGCTGAACAGATCGTAGT 221  
QY 160 CAAGATTACGATCAAGGTTATGTCACCTCTCATCGAGACCATTCATTACTATATG 219  
DB 222 CAATAATTACTACCAAGGTTATGTTACCTCAACCGTGACCATTCATTCTTACATGG 281  
QY 220 CAAGTTCCTTATGATGCATCATCAGTGAAGAGCTCCCTCATGAAAGATCCGAATATCA 279  
DB 282 GAAAGTTCCTTATGATGCGATTTATTAGTGAAGAGTTGTGTGATGCGATCCTTAATACCA 341  
QY 280 GTTGAAGGATTCAGACATTTGTCAATGAAATCAAGGGTGGTTATGTCAATTAAGTAACGG 339

Db 342 TTTTAAACAACTGACGGTTATCAATGAATCTTAGACGGTTACGTTATTAAAGTCAATGG 401  
QY 340 TAAATACCTATGTTACCTTAAAGGATGACGCTCATGCGGTAATATTCGACAAAGAAGA 399  
Db 402 CAACTATTATTGTTTACCTCAAGCAGGTAGTAGCGCAAAAACATTCGAACCAACAACA 461  
QY 400 GATTAACGTCAGAACGAGGAGCGAGTCAATATCAATCAATCAAA----- 443  
Db 462 AATTGCTGAGCAAGTAGTACCAAGGAACTAAAGAGAGCTAAAGAAAGGTTTATAGTCAAGT 521  
QY 444 -----GAGCAGATAATGCTGTTGCTGCAGCCAGAGCCCAAGGAGGTTA 486  
Db 522 GGCCCACTCAGTAAGAAGAGTTGCGGCAAGTCAATGAACCAAAAGACAGAGCGCTA 581  
QY 487 TACAACGGATGATGGGTATATCTTCAATGATCTGATATCAATTCAGGACACGGGTGATGC 546  
Db 582 TACTACAGAGTGGCTATATTTTAGTCCGACAGATATCATTTGATGATTTAGGAGATGC 641  
QY 547 TTATATCGTTCTCACGGGACCAATACCATTTCCCTTAAGATGATGTTATCAGCTAG 606  
Db 642 TTATTTAGTACCTCATGGTAATCACTATCATTTATTTCTTAAAAAGATTTGTCTCCAAG 701  
QY 607 CGAGTTAGTCTGCAGAGCCTATTGGAATGGGAGCAGGATCTGCTCCT 658  
Db 702 TGAGCTAGTCTGCAGACAGCCTACTGGAGTCAAAAACAGGTCGAGGTGCT 753

RESULT 38  
US-09-769-736-17  
; Sequence 17, Application US/09769736  
; Publication No. US20030138775A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbial Technics Limited  
; APPLICANT: Le Page, Richard WF  
; APPLICANT: Wells, Jeremy M  
; APPLICANT: Hanniffy, Sean B  
; TITLE OF INVENTION: Proteins  
; FILE REFERENCE: PWC/P21089wo  
; CURRENT APPLICATION NUMBER: US/09/769,736  
; CURRENT FILING DATE: 2003-02-14  
; PRIOR APPLICATION NUMBER: GB 9816335.5  
; PRIOR FILING DATE: 1998-07-27  
; PRIOR APPLICATION NUMBER: US 60/125163  
; PRIOR FILING DATE: 1999-03-19  
; NUMBER OF SEQ ID NOS: 212  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 2469  
; TYPE: DNA  
; ORGANISM: Streptococcus agalactiae  
US-09-769-736-17

Query Match 9.8%; Score 225; DB 10; Length 2469;  
Best Local Similarity 63.5%; Pred.No.6.1e-52;  
Matches 376; Conservative 0; Mismatches 183; Indels 33; Gaps 1;

QY 100 AAACCTTGACACAGATGAGTCAAGTAAGGAGGAGGATCAACGCCGCAACAAATGTTNAT 159  
Db 162 AAACAAAGATGGATCAATCAGTCTGAAAGGCACTCTGCTGAACAGATCTAGT 221  
QY 160 CAAGATTACGGATCAAGGTTATGTGACCTCTCATGAGACCAATATCATTTACTAATGG 219  
Db 222 CAAATTAATCTGACCAAGGTTATGTTACCTCACACGGTGACCAATATCAATTTTACCAATGG 281  
QY 220 CAAGGTTCTTTATGATGCAATCATCATGTAAGAGCTCTCATGAAGATCCGAATATCA 279  
Db 282 GAAAGTTCTTTATGATGCGATTTATTTAGTGAAGAGTTGTTGATGACGATCCTAATACCA 341  
QY 280 GTTGAAGGATTCAGACATTTGCTCAATGAATCAAGGTTGTTATGTTCAATTAAGGTAACGG 339  
Db 342 TTTTAAACAACTCAGAGCTTATCAATGAATCTTAGACGGTTATGTTATTAAAGTCAATGG 401  
QY 340 TAAATACCTATGTTACCTTAAAGGATGACGCTCATGCGGTAATATTCGACAAAGAAGA 399

Db 402 CAACTATTATTGTTTACCTCAAGCAGGTAGTAAGCGCAAAAACATTCGAACCAACAACA 461  
QY 400 GATTAACGTCAGAACGAGGAGCGAGTCAATATCAATCAATCAAA----- 443  
Db 462 AATTGCTGAGCAAGTAGTACCAAGGAACTAAAGAGAGCTAAAGAAAGGTTTATAGTCAAGT 521  
QY 444 -----GAGCAGATAATGCTGTTGCTGCAGCCAGAGCCCAAGGAGGTTA 486  
Db 522 GGCCCACTCAGTAAGAAGAGTTGCGGCAAGTCAATGAACCAAAAGACAGAGCGCTA 581  
QY 487 TACAACGGATGATGGGTATATCTTCAATGATCTGATATCAATTCAGGACACGGGTGATGC 546  
Db 582 TACTACAGAGTGGCTATATTTTAGTCCGACAGATATCATTTGATGATTTAGGAGATGC 641  
QY 547 TTATATCGTTCTCACGGGACCAATACCATTTCCCTTAAGATGATGTTATCAGCTAG 606  
Db 642 TTATTTAGTACCTCATGGTAATCACTATCATTTATTTCTTAAAAAGATTTGTCTCCAAG 701  
QY 607 CGAGTTAGTCTGCAGAGCCTATTGGAATGGGAGCAGGATCTGCTCCT 658  
Db 702 TGAGCTAGTCTGCAGACAGCCTACTGGAGTCAAAAACAGGTCGAGGTGCT 753

RESULT 39  
US-09-252-088-13  
; Sequence 13, Application US/09252088  
; Publication No. US20030031682A1  
; GENERAL INFORMATION:  
; APPLICANT: BRODEUR, Bernard R.  
; APPLICANT: RLOUX, Clment  
; APPLICANT: BOYER, Martine  
; APPLICANT: CHARLEBOIS, Isabelle  
; APPLICANT: HAMEL, Jose  
; APPLICANT: MARTIN, Denis  
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS  
; FILE REFERENCE: 8331-9002  
; CURRENT APPLICATION NUMBER: US/09/252,088  
; CURRENT FILING DATE: 1999-02-18  
; EARLIER APPLICATION NUMBER: US/60/075,425  
; EARLIER FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 5215  
; TYPE: DNA  
; ORGANISM: group B streptococcus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3)..(122)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (133)..(2511)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (367)..(2511)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: Complement ((2716)..(2946))  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: Complement ((2995)..(3252))  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: Complement ((3299)..(3676))  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: Complement ((3837)..(4124))  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: Complement ((4351)..(5214))  
US-09-252-088-13



```
; APPLICANT: Microbial Technics Limited
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21089wo
; CURRENT APPLICATION NUMBER: US/09/769,736
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: GB 9816335.5
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125163
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
US-09-769-736-71

Query Match          5.1%; Score 117; DB 10; Length 1455;
Best Local Similarity 54.8%; Pred. No. 1.1e-21;
Matches 334; Conservative 0; Mismatches 245; Indels 30; Gaps 4;

QY 1470 AACGTTTAGGAAACCAATGCGCAAATTTACCTACATGATGATGAGATTCAAGTAGCCA 1529
Db 257 AACGTAAAGGCAACCAATAGCCAGATTGTTCTACAGTGTCTCAAGAAATTGAAGAGGCAA 316

QY 1530 AGTTGGGAGGCAAGTACACACAGAGAGCGGTTTATCTTTGATCTCTGTGATATAACCA 1589
Db 317 AAAAAGCTGGTAAATACACCAATCTGATGTTTGTACATTTTGTGCTAAAGATATTAAAA 376

QY 1590 GTGATGAGGGGATGCTTATGTAATCCACATATGACCCATAGCCACTGGATTAAAAAAG 1649
Db 377 AAGATACAGGTACAGGTATGTCTATCCATATGACATGATGATGCTGGTACCAAGA 436

QY 1650 ATAGTTTCTGTAAGCTGAGAGAGCGGAGCCAGGCTTATGCTAAAGAGAAAGTTTGA 1709
Db 437 AAGATTTATCAGAGTCGGAATTTAAAGCAGCTCAAGAAATTTCTTTCAGGAAAA----- 489

QY 1710 CCCCTCTTCGACACACCTACAGGATTCAGGAATCTAGGAAATCTAGGCAAAAGAGCAGCTA 1769
Db 490 -----TCTGAAGCAATCAAGACAAACCAACCAAC-----AGTAAACAGCTCAAGAAA 538

QY 1770 TCTACACCGCTGGAAGCAGCTAAGAGGTGCCACTTGTATGCTATGCTTACAACTTTC 1829
Db 539 TCTATGAGCAATTAACCAAGCAATTTGTTAACTGAGATTTATTTTGGAAATTG 598

QY 1830 AATATACTGTAGAAGTCAAAAACGGTAGTTTAACTATCATCTCATATGACCATTAACATA 1889
Db 599 CACAAGCGACACACTATAAGATGGTACATTTGTAATTCCTCATATAAGATCATTACCAT 658

QY 1890 ACATCAAAATTTGAGTGGTTTGAGAAAGGCGCTTTA-----TGAGGACCTTAAGGGGT 1940
Db 659 ATGTGGAATTAATAATGTTTGTATGAAGAAAAAGATCTTTTAGCTGATTTCAGATAAGACAT 718

QY 1941 ATACTCTTGAGATCTTTTGGGACCTGTCAGTACTATGTCGAACATCCAAACGACGTC 2000
Db 719 ATTCTTTAGAAGACTATTATTAGCTACGGCTAAATATTACATGATGACCCAGAAAAACGTC 778

QY 2001 CGCATTCAGATATGTTTGGTAAAGCTAGCGACCATGTTTCAAGAAACAAAAATGGTC 2060
Db 779 CTAA---AGTTGAGAGTGGGTAAAGATGCTGAATTTTATAGGAAAGGACTCTTAATA 835

QY 2061 AAGCTGATA 2069
Db 836 AAGCAGATA 844

RESULT 42
US-10-087-192-463/c
; Sequence 463, Application US/10087192
; Publication No. US20020182586A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 463
; LENGTH: 31124
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(31124)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-463

Query Match          2.5%; Score 58.2; DB 13; Length 31124;
Best Local Similarity 52.2%; Pred. No. 0.00032;
Matches 129; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 2043 AAAGAAACAAAAATGGTCAAGCTGTATACCAATCAACGAAAAAACCAAGCCGAGGAGAAAC 2102
Db 18657 ACAAGACCAAAAGAAAAGAACCAAGACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 18598

QY 2103 CTCAGACAGAAAAACCTTGAGGAAGAAAAACCCCTCCAGAAAGAGAAAAACCGCAAGCGGAAAC 2162
Db 18597 AAGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 18538

QY 2163 CAGAGTCTCCAAACCAACAGAGAGAACCCAGAGGATCACCAGAGGAATCAGAAGAAACCTC 2222
Db 18537 AAGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 18478

QY 2223 AGTCCGAGACTGAAAAAGGTTGAAGAAAAAATCGAGAGAGGCTGAAGATTTACTTGAAGAAA 2282
Db 18477 GAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 18418

QY 2283 TCCAGGA 2289
Db 18417 AAGAAGA 18411

RESULT 43
US-09-957-956-6
; Sequence 6, Application US/09957956
; Publication No. US20030130215A1
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; TITLE OF INVENTION: Genomic Polynucleotide Fragments From Chromosome 7
; FILE REFERENCE: JR-14, 000-US
; CURRENT APPLICATION NUMBER: US/09/957,956
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/234,422
; PRIOR FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 45980
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-957-956-6

Query Match          2.4%; Score 54.8; DB 10; Length 45980;
Best Local Similarity 52.7%; Pred. No. 0.0038;
Matches 119; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
```











GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

M protein - nucleic search, using frame\_plus\_p2n model

run on: May 1, 2004, 04:16:59 ; Search time 134 Seconds  
(without alignments)  
3159.907 Million cell updates/sec

title: US-09-765-272A-66  
effect score: 4019  
sequence: 1 CSVELGHRGQGVKKESNRV.....TEKVEKLRREADLLGRIQD 763

coring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

searched: 682709 seqs, 277475446 residues

total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Listing first 150 summaries

post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

command line parameters:  
MODEL=frame+ p2n.model -DEV=xlh  
Q=/cgn2\_1/USPTO.spool/US09765272/runat 30042004\_081051\_11606/app\_query.fasta\_1.903  
DB=Issued Patents NA -QWMT=fastac -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPFCI=0  
LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -WATRIX=blomsum62 -TRANS=human40.cdi  
LIST=150 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=50  
MODE=LOCAL -OUTFMT=prc -NORM=exnat 30042004\_081051\_11606 -NCPU=6 -ICPU=3  
USBR=US09765272@cgn\_1\_156@runat 30042004\_081051\_11606 -LONGLOG  
NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:  
5: /cgn2\_6/ptodata/2/ina/PCUS\_COMB.seq:  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result	No.	Score	Query Match Length DB ID Description
1	4013	99.9	2290 3 US-08-961-083-65 Sequence 65, Appl
2	4013	99.9	2291 4 US-09-536-784-65 Sequence 11, Appl
3	3919	97.5	2531 4 US-09-468-656A-11 Sequence 5, Appl
4	3644.5	90.7	2531 4 US-09-468-656A-5 Sequence 94, Appl
5	3641.5	90.6	8195 4 US-08-961-527-94 Sequence 9, Appl
6	2658.5	66.1	2451 4 US-09-468-656A-9 Sequence 55, Appl
7	2649.5	65.9	2389 3 US-08-961-083-55 Sequence 55, Appl
8	2649.5	65.9	2389 3 US-09-536-784-55 Sequence 243, App
9	2424	60.3	2359 4 US-08-961-527-243 Sequence 355, App
10	1377	34.3	973 4 US-08-961-527-355 Sequence 192, App
c 11	1202.5	29.9	6867 4 US-08-961-527-192 Sequence 258, App
12	1200	29.9	1684 4 US-08-961-527-258

13	1167.5	29.0	1455	4	US-09-468-656A-7	Sequence 7, Appl
14	1141.5	28.4	1342	3	US-08-961-083-181	Sequence 181, App
15	1141.5	28.4	1342	4	US-09-536-784-181	Sequence 181, App
c 16	512	12.7	841	2	US-08-743-637B-34	Sequence 34, Appl
c 17	512	12.7	841	2	US-08-526-840B-34	Sequence 34, Appl
18	177	4.4	3095	6	5231168-1	Patent No. 5231168
19	165.5	4.1	2232	1	US-08-257-073-4	Sequence 4, Appl
c 20	165	4.1	2392	4	US-09-216-393B-331	Sequence 331, App
21	163	4.1	2393	4	US-09-216-393B-329	Sequence 329, App
22	163	4.1	4200	1	US-08-242-932-1	Sequence 1, Appl
23	163	4.1	4200	1	US-08-714-481-1	Sequence 1, Appl
24	163	4.1	4200	3	US-08-923-992A-1	Sequence 1, Appl
25	163	4.1	4200	5	PCT-US95-06111-1	Sequence 1, Appl
c 26	160.5	4.0	4522	4	US-08-956-171B-39	Sequence 39, Appl
c 27	159.5	4.0	80161	3	US-09-036-987A-1	Sequence 1, Appl
c 28	159.5	4.0	80161	3	US-09-370-700-1	Sequence 1, Appl
c 29	159.5	4.0	80161	4	US-09-603-207-1	Sequence 1, Appl
30	158	3.9	1278	4	US-09-286-981B-29	Sequence 29, Appl
31	157.5	3.9	1272	4	US-09-286-981B-31	Sequence 31, Appl
32	156.5	3.9	1276	4	US-09-286-981B-30	Sequence 30, Appl
33	156.5	3.9	12665	4	US-08-961-527-134	Sequence 134, App
c 34	154.5	3.8	28473	4	US-08-961-527-83	Sequence 83, Appl
c 35	153	3.8	580073	4	US-08-545-528D-1	Sequence 1, Appl
36	151	3.8	3384	3	US-08-923-992A-5	Sequence 5, Appl
37	151	3.8	10711	4	US-08-961-527-145	Sequence 145, App
38	150.5	3.7	1338	4	US-09-286-981B-26	Sequence 26, Appl
39	150.5	3.7	5468	4	US-09-535-008-66	Sequence 66, Appl
40	150.5	3.7	5471	4	US-09-535-008-1	Sequence 1, Appl
41	150.5	3.7	5477	4	US-09-535-008-74	Sequence 74, Appl
42	150.5	3.7	5480	4	US-09-535-008-70	Sequence 70, Appl
43	150.5	3.7	5564	4	US-09-535-008-68	Sequence 68, Appl
44	150.5	3.7	5567	4	US-09-535-008-64	Sequence 64, Appl
45	150.5	3.7	5573	4	US-09-535-008-76	Sequence 76, Appl
46	150.5	3.7	5576	4	US-09-535-008-72	Sequence 72, Appl
47	150	3.7	3279	3	US-08-446-137B-1	Sequence 1, Appl
48	149.5	3.7	5471	4	US-09-535-008-62	Sequence 62, Appl
49	149.5	3.7	4403765	3	US-09-103-840A-2	Sequence 2, Appl
50	149.5	3.7	3294	3	US-09-103-840A-1	Sequence 1, Appl
51	149	3.7	3492	3	US-08-923-992A-9	Sequence 9, Appl
52	149	3.7	3492	3	US-08-923-992A-7	Sequence 27, Appl
53	147	3.7	5295	1	US-09-221-017B-27	Sequence 27, Appl
54	146.5	3.6	6605	1	US-08-769-309A-4	Sequence 4, Appl
55	146.5	3.6	6605	3	US-08-994-570-4	Sequence 4, Appl
56	146.5	3.6	6608	4	US-09-220-132-58	Sequence 58, Appl
57	145	3.6	2370	4	US-09-134-000C-1534	Sequence 1534, Ap
58	145	3.6	3711	4	US-09-566-921-126	Sequence 126, App
59	144	3.6	3312	3	US-08-923-992A-3	Sequence 3, Appl
60	143.5	3.6	1338	4	US-09-286-981B-23	Sequence 23, Appl
61	143	3.6	5710	1	US-08-425-061-6	Sequence 6, Appl
62	143	3.6	5710	1	US-08-825-886-6	Sequence 6, Appl
63	143	3.6	5710	4	US-08-989-890-6	Sequence 6, Appl
64	143	3.6	269223	4	US-09-596-002-41	Sequence 41, Appl
65	142.5	3.5	3534	4	US-09-134-001C-2269	Sequence 2269, Ap
66	142	3.5	31096	4	US-08-956-371E-59	Sequence 59, Appl
67	141.5	3.5	3127	4	US-09-620-312D-613	Sequence 61, App
68	141	3.5	9408	4	US-09-418-710-14	Sequence 14, Appl
69	140.5	3.5	3546	4	US-09-540-236-1884	Sequence 1884, Ap
70	140.5	3.5	6744	1	US-08-119-125A-2	Sequence 2, Appl
71	140	3.5	3752	4	US-08-961-527-208	Sequence 208, App
72	140	3.5	30549	4	US-09-134-001C-322	Sequence 322, App
73	139	3.5	4118	1	US-08-119-125A-3	Sequence 3, Appl
74	139	3.5	5008	3	US-09-110-517-1	Sequence 1, Appl
75	139	3.5	14770	4	US-09-220-132-30	Sequence 30, Appl
76	138.5	3.4	3489	2	US-08-728-323A-1	Sequence 1, Appl
77	138.5	3.4	3489	4	US-09-238-568-1	Sequence 1, Appl
78	138.5	3.4	3489	4	US-09-410-399-1	Sequence 1, Appl
79	138.5	3.4	32207	2	US-08-770-379-20	Sequence 20, Appl
c 80	138.5	3.4	32207	3	US-08-757-669A-20	Sequence 20, Appl
c 81	138.5	3.4	32207	3	US-09-230-371A-20	Sequence 20, Appl
82	138	3.4	11907	3	US-08-061-376-4	Sequence 4, Appl
83	138	3.4	14255	1	US-08-327-359-1	Sequence 1, Appl
84	138	3.4	14255	1	US-08-327-352-1	Sequence 1, Appl
85	138	3.4	14255	1	US-08-306-691B-55	Sequence 55, Appl





COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

**MEDIUM TYPE:** Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/536,784

AFFIDAVIT NUMBER: 03/03/  
 FILING DATE: 30-Oct-1997

CLASSIFICATION: <Unknown>  
 FILING DATE: 30-OCT-1997

CLASSIFICATION: ~~CONFIDENTIAL~~  
PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/961 083

APPLICATION NUMBER: 08/961  
FILING DATE: OCT-30-1997

FILING DATE: OCT-30-1997  
ATTORNEY / AGENT INFORMATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Michallo E Marks

NAME: Michelle S. Marks  
REGISTRATION NUMBER: 47 05

REGISTRATION NUMBER: 41,977

REFERENCE/DOCKET NUMBER: E

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

FORMATION FOR SEQ ID NO: 65:

### SEQUENCE CHARACTERISTICS:

LENGTH: 2290 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO

Alignment Scores:

Argument scores:	
pred. No.:	0
Length:	2290

PLD: NO.:	0	LENGCH:	2230
SCORE:	4013.00	MATCHES:	763

SCORE:	4013.00	MATCHES:	763
Percent similarity:	100.00%	Conservative:	0

Percent similarity:	100.00%	Conservative:	0
Best local similarity:	100.00%	Mismatches:	0

```

BEST LOCAL SIMILARITY: 100.00%
Mismatched: 0
Query Match: 99.85%
Index Match: 99.85%
Mismatched: 0

```

Query Match:	99.85%	Indels:	0
DB:	A	Gaps:	0

UIS-09-765-272A-66 (1-763) x UIS-09-536-784-65 (1-2290)

1	CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal	20
2	TGTTCTCTATGAACCTGGTCTGTCACCAAGCTGGTCAGGTTACAAAGAGTCTAATCGAGTT	61
21	SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal	40
62	TCTTATATAGATGGTGTATCAGCGTGGTCAAAAGGCAGAAACCTCACCACGATGAAGTC	121
41	SerLysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyr	60
122	AGTAAGAGGGGGGATCAACGCCGACAAATNGTNATCAAGATTACGGATCAAGGTTAT	181
61	ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle	80
182	GTGACCTCTCATGGAGACCATTTATCATTTACATAATGGCAAGGTTCCITATGATGCCATC	241
81	IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal	100
242	ATCAGTGAAGAGCTCCTCATGAAAGATCCGATTTATCATGTTCAAGGATTCAGACATTGTC	301
101	AsnGluIleLysGlyGlyTyr*ValIleLysValAsnGlyLysTyrValTyrIleLys	120
302	AATGAATATCAAGGGTGGTTATGTCTAATTAAGSTAAACGGTAAATCATATGNTACCTTAAG	361
121	AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu	140
362	GATCGAGCTCATGGGATTAATATTCGGACAAAGAAGAGAGATTAACTCAGAAGCAGGAA	421
141	ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly	160
422	CGCAGTCATATCATCACTCAAGAGCAGATATGCTGTGTGCTGCAGCCAGAGCCCAAGGA	481
161	ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly	180
482	CGTTATACAACGGATGATGGGTATATCTTCATGTCATCTGATCATTCATTGAGACACGGGT	541





```
1018 CGTTATCGTTCAACATTGGGTACCAGATTCCAGACCAGAACCAAGTCCACAAACCG 1077
341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360
1078 ACTCCAGAACCTAGTCCAGT-----CCGCAACACAGCTCCAGCAAT 1119
361 ProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380
1120 CCAATTGATGGAAATTGCTCAAGAAAGCTGTTCCGAAAGTAGCGGATGTTATGCTTT 1179
381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400
1180 GAGGAGAATGGAGTTCTCGTTATATCCCGCCGCAAGGATCTTTCCAGCAAAAACAGCAGCA 1239
401 GlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLys 420
1240 GGCATTGATAGCAAACTGCCAAGCAGGAAAGTTTATCTCATAGCTAGGAACCTAGAAA 1299
421 ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440
1300 ACTGACCTCCCATCTAGTATGATCAGAAATTTTCAATAAAGCTTATGACTTACTTAGCAAGA 1359
441 IleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460
1360 ATTCACCAAGATTACTTGATATTAAGSTCGCAAGTTGATTTTGAGGCTTTGGATAAC 1419
461 LeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeu 480
1420 CTGTTGGAACGACTCAAGATGCTCAAGTGAATAGTCAAGTAGTGGAAAGATATCTT 1479
481 AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr 500
1480 GCCTTCTTAGTCCGATTCGTCATCCAGACGTTTGGAAAAACCAATGCGCAATTAACC 1539
501 TyrThrAspGluLeuGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGly 520
1540 TACACTGATGATGAGATTCAAGTAGGCAAGTTCGCGAGCAAGTACACAGCAGAACGCGT 1599
521 TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis 540
1600 TATATCTTTGATCCTCGTGATATAACCAAGTAGTGGGGGATGCTTATGTAATCTCCACAT 1659
541 MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAla 560
1660 ATGACCCATAGCCACTGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGCGAGCC 1719
561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580
1720 CAGGCTTATGCTTAAGAGAAAGGTTTGACCCCTCCTTCGACAGACCATCAGGATTCAGGA 1779
581 AsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysVal 600
1780 AATACCTGAGCAAAAGAGCAGAGAGCTATCTTACAAACCGMGTTGAAGCAGCTTAAGAGGTG 1839
601 ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu 620
1840 CCATTGATCGTATGCTTACATCTTCAATATCTAGTACTAGAGTCAAAACGGTAGTTTA 1899
621 IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeu 640
1900 ATCATACCTCATTATGACCATACCAATAACATCAATCAATTTGAGTGGTTTACGAAGGCGCTT 1959
641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrVal 660
1960 TATGAGCACCTTAAGGGGTATACTCTTTGAGGATCTTTTGGCGACTGTCAAGTACTATGTC 2019
661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680
2020 GAACATCCAAACGAACTCCGCATTCAGATTAATGGTTTGGTAACGCTAGGACCACTGTT 2079
681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLys 700
```

```
Db 2080 CAAAGAAACAAAATGTCAGCTGATACCAATCAACCGAAAAACCAAGCGAGGAGAA 2139
Qy 701 ProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLys 720
Db 2140 CCTCAGACAGAAAAACCTGAGAAAGAAACCCCTCGAGAAGAGAAACCGCAAGCGAGAA 2199
Qy 721 ProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro 740
Db 2200 CCAGAGTCTCCAAACCAACACAGAGGAACCAAGAATCACCAGAGGAATCAGAAAGACCT 2259
Qy 741 GlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLys 760
Db 2260 CAGTCCAGACTGAAAAGTTGAAGAAAACTGAGAGAGGCTGAAGATTACTTGGAAAA 2319
Qy 761 IleGlnAsp 763
Db 2320 ATCCAGGAT 2328
RESULT 4
US-09-468-656A-5
; Sequence 5, Application US/09468656A
; Patent No. 6582706
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Group A and B Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-444
; CURRENT APPLICATION NUMBER: US/09/468,656A
; CURRENT FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-468-656A-5
Alignment Scores:
Pred. No.: 0 Length: 2531
Score: 3644.50 Matches: 698
Percent Similarity: 92.41% Conservative: 20
Best Local Similarity: 89.83% Mismatches: 40
Query Match: 90.68% Indels: 19
DB: 4 Gaps: 4
US-09-765-272A-66 (1-763) x US-09-468-656A-5 (1-2531)
Qy 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 58 TGTTCTATGAACTTGGTCGTCACCAAGCTGGTCAGTTAGAAAGTCTATTCGAGTT 117
Qy 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 TCTTATATAGATGGTATCAGCTGGTCAAAAGGCAGAAAACTTCACACCATCAAGTCAAGTTC 177
Qy 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGTAAGAGGAGGGGATCAACGCCCAACAAATCGTCATCAAGATTACGGATCAAGGTTAT 237
Qy 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTGACCTCTCATGGAGACCATTTATCTATATAATGGCAAGGTCCCTTATGATGCATC 297
Qy 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 298 ATCAGTGAAGAGCTCTCTCATGAAGATCCGATTTATCAGTTCAAGGATTCAGACATTGTC 357
Qy 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
```

358 AATGAAATCAAGGGTGGTTATCTTATCAAGGTAGATGCGAAATACTATCTTACCTTAAG 417  
121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140  
418 GATGAGCTCATCGGATTAATATTGCGACAAAGAGAGATTAAAGCTCAGAGCAGGAA 477  
141 ArgSerHisAsnHisAsnSerArgAla---AspAsnAlaValAlaAlaAlaAlaGln 159  
478 CACAGTCATAATCACGGGGTGGTTCTTAAGCATCAAGCAGTAGTTGTCAGCCAGAGCCAA 537  
160 GlyArgThrThrAspAspGlyThrIlePheAsnAlaSerAspIleIleGluAspThr 179  
538 GGACGCTATACAAACGATGATGGTTATATCTTCAATGCAATCTGATATCAATTGAGGACAG 597  
180 GlyAspAlaThrIleValProHisGlyAspHisThrHisThrIleProLysAsnGluLeu 199  
598 GGTGATGCTTATCTGCTTCCACGGCGACCATACCATACATTCCTAAGAATGAGTTA 657  
200 SerAlaSerGluLeuAlaAlaAlaGluAlaThrTrpAsnGlyLysGlnGlySerArgPro 219  
658 TCAGCTAGCGAGTAGCTGCTGCAGAGCGCTATTGGAATGGGAAGCAGGATCTCGTCT 717  
220 SerSerSerSerSerTrpAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsn 239  
718 TCTTCAAGTTCTAGTTATAATGCAATCCAGCTCAACCAAGATTGTACAGAGAACCAAT 777  
240 LeuThrValThrProThrThrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArg 259  
778 CTGAGCTGTCACCTCAACTTATCATCAAAATCAAGGGGAACAAATTCAGGCCCTTTACGT 837  
260 GluLeuThrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAsp 279  
838 GAATTTGTATGCTAAACCTTATCAGAACGCCATGTGGAATCTGATGGCTTATTTTCGAC 897  
280 ProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisThr 299  
898 CCAGCGCAATCACAAATCGAACCGCGCAGAGGTGTAGCTGTCCCTCATGTGAACCAATTAC 957  
300 HisPheIleProThrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIlePro 319  
958 CACTTTATCCCTTATCAACAAATGCTGAATTTGGAACCAAGAAATGCTGCTATTTATCC 1017  
320 LeuArgThrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerProGln 339  
1018 CTTCGTATTCGTTCAACCATTTGGTGACAGATTCAAGACAGAACCAACAGTCCACAA 1077  
340 SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer 359  
1078 TCGACTCCGGAACCTAGTCCAAAGTCGCAACCTGCACCAATCTCAACCCAGCTCCAAAGC 1137  
360 AsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyThrVal 379  
1138 AATCCAATTGATGAGAAATGGTCAAGAGAGCTGTTCGAAAGATGACGATGTTATGTC 1197  
380 PheGluGluAsnGlyValSerArgThrIleProAlaLysAspLeuSerAlaGluThrAla 399  
1198 TTGAGGAGAAATGGAGTTCTCTGTTATATCCCAAGCAAGATCTTTCAGCAGAAACAGCA 1257  
400 AlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys 419  
1258 GCAGCAATTGATAGCAAACTGGCCAGCAGCAGGAAGTTTATCTATAAGCTAGGAGCTAAG 1317  
420 LysThrAspLeuProSerSerAspArgGluPheThrAsnLysAlaThrAspLeuLeuAla 439  
1318 AAAAAGTACCTCCATCTAGTATCGAGAAATTTTACAATTAAGGCTTATGACTTACTAGCA 1377  
440 ArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAsp 459  
1378 AGAATTCACCAAGATTACTTGTATATAAAGGTCGACAGTTGATTTTGGGCTTTGGAT 1437  
460 AsnLeuLeuGluArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIle 479  
1438 AACCTGTTGGACGATCTCAAGGATGTCCCAAGTGATTAAGTCAAGTTAGTGGATGATTT 1497

QY 480 LeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIle 499  
Db 1498 CTTGCCCTTCTAGCTCGATTGCTCATCCAGAAAGTTAGGAAACCAAAATCGCAAAAT 1557  
QY 500 ThrThrThrAspAspGluIleGlnValAlaLysIleAlaGlyLysThrThrGluAsp 519  
Db 1558 ACCTACACTCATGATGAGATTCAAGTAGCCAGGTTGGCAGCAAGTACACACAGAAAGC 1617  
QY 520 GlyThrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaThrValThrPro 539  
Db 1618 GGTATATCTTTGATCTCTGATATATACCACTGATGAGGGGATGCCTATGTAATCTCA 1677  
QY 540 HisMetThrHisSerHisThrIleLysLysAspSerLeuSerGluAlaGluAlaAla 559  
Db 1678 CATATGACCCATAGCCACTGGATTAAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGCA 1737  
QY 560 AlaGlnAlaThrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSer 579  
Db 1738 GCCCAGGCTTATGCTAAAGAGAAAGGTTTGACCCCTCTTCACAGACCATCAGGATTCA 1797  
QY 580 GlyAsnThrGluAlaLysGlyAlaGluAlaIleThrAsnArgValLysAlaAlaLysLys 599  
Db 1798 GGAATATCTGAGGCAAAAGGAGAGAGCTATCTACAAACCGGTGAAAGCAGCTAAGAG 1857  
QY 600 ValProLeuAspArgMetProTrpAsnLeuGlnThrValGluValLysAsnGlySer 619  
Db 1858 GTGCCACTTATGATCGTATGCTTACAACTTCAATATCTAGTAAGTCAAAACCGTAGT 1917  
QY 620 LeuIleIleProHisThrAspHisThrHisAsnIleLysPheGluTrpPheAspGluGly 639  
Db 1918 TTAATCATACCTCATTTATGACCAATTACCAATACATCAAAATTTAGTGTGTTTACGAAAGC 1977  
QY 640 LeuThrGluAlaProLysGlyThrThrLeuGluAspLeuAlaThrValLysThrThr 659  
Db 1978 CTTTATCAGGCACCTTAAGGGGTATCTCTTGGAGATCTTTTGGCGACTGTCAAGTACTAT 2037  
QY 660 ValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHis 679  
Db 2038 GTCCAAATCCAAACGAACGTCCGCAATTCAGATATGTTTGGTAAACGCTAGCAGCAT 2097  
QY 680 ValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGlu 699  
Db 2098 GTTCGTAATAATAAG-----GTAGACCAAGACAGTAAACCTGATGAAGAT 2142  
QY 700 LysProGlnThrGluLysProGluGluThrProArgGluGluLysProGlnSerGlu 719  
Db 2143 AAGAAACATGATGAAGTAAGTAGGACCAACTCACCCCTGAATCTGTAGAAAGAGAGATCAC 2202  
QY 720 LysProGluSerPro-----LysProThrGluGluProGluGluSer 733  
Db 2203 GCTGTTTAATCTTTCAGCAGATAATCTTTATAAACCAAGCAGCTGATACGAGAGACA 2262  
QY 734 ProGluGluSerGluGlu-----ProGlnValGluThrGluLys 746  
Db 2263 GAGCAAGAGCTGAAGTACCACAGATCAGGCTGAAATTCCTCAAGTAGAGAAATCTGTT 2322  
QY 747 ValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763  
Db 2323 ATTAACGTAAGTAGCAGATGCGGAGGCTTGTAGAAAGAAAGTAACAGAT 2373

## RESULT 5

US-08-961-527-94

; Sequence 94, Application US/08961527

; Patent No. 6420135

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville





286 ArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGlu 305  
955 CGAAACAGTAGAGGTGTTGCAAGTGCACACGAGATCATTAACCACTTCATCCCTTACTCT 1014  
306 GlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrArgSerIle 325  
1015 CAATGTCGTGAATTGGAAGAAGCAATCGTGTATTTCCCTTCCTGTTATCGTTCAAA 1074  
326 HisTyrValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSer 345  
1075 CATTCGGGTACAGATTCAAGGCCAGAAACAACAAGTCCACACCGACTCCGGAACTAGT 1134  
346 ProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLys 365  
1135 CCAGCCCGCAACCTGCACCAATCTTAAATATAGACTCAAAAT-----TCTTCT 1182  
366 LeuValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyVal 385  
1183 TTGGTTAGTCAGTGTGTAAGAAAGTGGGGAAGGATATGTTATCGAAGAAAGGCGATC 1242  
386 SerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys 405  
1243 TCTGTTATGTTCTTGGCAAGATTACCACTGAACTGTAAATCTTGAAGCAAG 1302  
406 LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSer 425  
1303 TTATCAAAACAAGAGAGTGTTCACACACTTTAACTGCTTAAAAAAGAAATGTTGCTCT 1362  
426 SerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeu 445  
1363 CGTGACCAAGAATTTATGATAAGCATATATCTGTTAACTGAGGCTCATAAAGCCTTG 1422  
446 LeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeu 465  
1423 TTTGNAANTAGGGTCGTAATCTGATTTCCAGCTTAGACAATATTATAGAGCTTG 1482  
466 LysAspVal\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeuAlaPheLeuAlaPro 485  
1483 AATGATGATGACTAATAAAGAAAATGTTAGATGATTATTGTCATTCCTAGCACCA 1542  
486 IleArgHisProGluArgLeuGlyLysProAlaGlnIleThrTyrThrAspAspGlu 505  
1543 ATTACCATCCAGACGCACTTGGCAACAAATCTCAAAATGAGTATACTGAAGACGAA 1602  
506 IleGlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspPro 525  
1603 GTTCGTTATGCTCAATTAGCTGATAGTATACACGTCAGATGGTTTACATTTTGTGAA 1662  
526 ArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHis 545  
1663 CATGATATAATCAGTGATGAAGGAGATGCATATGTAACGCTCATATGGCCATAGTCAC 1722  
546 TrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAlaLys 565  
1723 TGGATTGAAAAGATAGCCTTCTGATGAAGAAAAAGTTGCACTCAAGCCCTATATAA 1782  
566 GluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLys 585  
1783 GAAAAGGATATCTTACCTCCATCTCCAGACGAGATGTTAAAGCAAAATCCAACTGGAGAT 1842  
586 GlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArgMet 605  
1843 AGTCAGCAGCTATTATCAATCGTGTGAAAGGGAAGAAAGAAATCCACTCGTTCGACTT 1902  
606 ProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleIleProHisTyr 625  
1903 CCATATATGTTGAGCATACAGTTCAGGTTAAAAACGGTAATTTGATTATTCTCATAG 1962  
626 AspHisTyrHisAsnIleLysPheGluThrPheAspGluGlyLeuTyrGluAlaProLys 645  
1963 GATCATTAACCAATAATTATTAATTTGTTGGTTTGTATGATCACACATACAAAGCTCCAAAT 2022

646 GlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGlu 665  
2023 GGCTATACCTTGAAGATTTGTTGGACGATTAAGTACTACGTAGAACACCCCTGACGAA 2082  
666 ArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsn 685  
2083 CGTCCACATCTTAATGATGGGCAATCCAGTCAGCATGTGTAGGCAAGAAAGAC 2142  
686 GlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLys 705  
2143 CACAGTGAAGATCCAAATAAGAACTTCAAAGCGGATGAAGAG----- 2184  
706 ProGluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSerProLys 725  
2185 CCAGTAGAGGAAACA----- 2199  
726 ProThrGluGluProGluGluSerProGluGluSerGluGluProGlnValGluThrGlu 745  
2200 -----CTGCTGAGCCAGAGTCCCTCAAGTAGAGACTGAA 2235  
746 LysValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763  
2236 AAAGTAGAAGCCCACTCAAAGAGCAGAGATTTTGCTTCGAAAGTAACGAT 2289

RESULT 7  
US-08-961-083-55  
; Sequence 55, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,083  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 55:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2389 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-083-55

Alignment Scores:  
Pred. No.: 5,96e-243 Length: 2389  
Score: 2649.50 Matches: 519  
Percent Similarity: 76.71% Conservative: 77  
Best Local Similarity: 66.80% Mismatches: 132  
Query Match: 65.92% Indels: 49  
DB: 3 Gaps: 8

US-09-765-272a-66 (1-763) x US-08-961-083-55 (1-2389)

1019 TGGGTACAGATTCAAGCCGAGAAACAACCAAGTCCACCAACCGACTCCGGAACCTAGTCCA 1078  
QY SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366  
Db GCGCGCAACCTGCACCAAACTTTAAA--ATAGACTCAAT-----TCCTTTCTTG 1126  
QY ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386  
Db GGTAGTCAGCTGGTACGAAAGAGTTGGGGAAGGATATGATTATCGAAGAAAGGCGCATCTCT 1186  
QY ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406  
Db CGTTATGCTTTTGGCAAGAGTTTACCCTGAACTGTTTAAAAATCTTGAAGCAAGTTA 1246  
QY AlaLysGlnGluSerLysHisLysLysLeuGlyAlaLysLysLysThrAspLeuProSerSer 426  
Db TCAAAACAAGAGAGTGTTCACACACTTTAACTGCTTAAAAAAGAAATGTTGCTCTCTCGT 1306  
QY AspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeu 446  
Db GACCAAGAAATTTATGATAAGCATATAATCTGTAACTGAGGCTCATAAAGCCTTGTTT 1366  
QY AsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLys 466  
Db GNAATAAGGGTCGTAATTTCTGATTTCCAAAGCTTAGACAAATTTATTAGAACGCTTGAAT 1426  
QY AspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeuAlaPheLeuAlaProIle 486  
Db GATGATCGACTAATAAGAAATAATTTGGTAGATGATTATTTGGCATCTCTAGCACAAT 1486  
QY ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIle 506  
Db ACCATCCAGAGCGACTGGCAACCAATTTCTCAATTTGAGTATCTAGAGAGAGTT 1546  
QY GlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspProArg 526  
Db CGTATTTGCTCAATTAGCTGATAAGTATATCAACGTCAGATGGTTACATTTTGTATGAACAT 1606  
QY AspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTyr 546  
Db GATATAATCAGTGATGAAGGAGATCATATGTAACGCTCATATGGCCATAGTCACTGG 1666  
QY IleLysLysAspSerLeuSerGluAlaGlyArgAlaAlaAlaGlnAlaTyrAlaLysGlu 566  
Db ATTGCAAAAGATAGCTTTCTGATAGGAAAGTTGGAGCTCAAGCTTATCTATAAGAA 1726  
QY LysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGly 586  
Db AAAGTATCTTACCTCCATCTCCAGAGCGAGATGTTTAAAGCAATCCAACTGGAGATAGT 1786  
QY AlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArgMetPro 606  
Db GCAGCAGCTATTATCAATCGTGTGAAGGGGAAAAACGAATTCACCTGCTTCGACTTCCA 1846  
QY TyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleIleProHisTyrAsp 626  
Db TATATGTTGAGTACATAGTTGAGTTAAACGGTAATTTGATTTATTTCTCTATAAGAT 1906  
QY HisTyrHisAsnIleLysPheGluThrPheAspGluGlyLeuTyrGluAlaProLysGly 646  
Db CATTAACATAATATTAATTTGCTTTGTTGATGATCACACATACAAAGCTCCAAATGGC 1966  
QY TyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValGluHisProLeuGluArg 666  
Db TATACCTTGAAGATTTCTTTGCGAGCATTAAGTACTACGTAGAACACACCTCGCAACGT 2026  
QY ProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGly 686  
Db CCACATCTTAATGATGATGGGCAATCCAGTACGATGTTGTTAGGCAAGAAAGACCCAC 2086  
QY GlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLysPro 706  
Db AGTGAAGATCCAAATAAGAACTTCAAGCGGATGAAGAG-----CCCA 2128



QY 707 GluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSerProLysPro 726  
Db 2129 GTAGAGGAAACA-----2140  
QY 727 ThrGluGluProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLys 746  
Db 2141 -----CCGCGTAGCCAGAGGCCCTCAAGTAGAGCTGAAAAA 2179  
QY 747 ValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763  
Db 2180 GTAGAGCCCACTCAAGAGCAGAGAGTTTGTCTTGCGAAGTAACGGAT 2230

RESULT 8  
US-09-536-784-55  
Sequence 55, Application US/09536784  
Patent No. 6573082  
GENERAL INFORMATION:  
APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/536,784  
FILING DATE: 30-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/961,083  
FILING DATE: OCT-30-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PB34093  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 55:  
US-09-536-784-55

Alignment Scores:  
Pred. No.: 5.96e-243 Length: 2389  
Score: 2649.50 Matches: 519  
Percent Similarity: 76.71% Conservative: 77  
Best Local Similarity: 66.80% Mismatches: 132  
Query Match: 65.92% Indels: 49  
DB: 4 Gaps: 8

US-09-765-272A-66 (1-763) x US-09-536-784-55 (1-2389)  
QY 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21  
Db 2 TCCTACGGTTGGGACTGTATCAAGCTAGAACGGTT---AAGGAAATAAATCGTGTTC 58

QY 22 TyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluValSer 41  
Db 59 TATATAGTGAACAAAGCAGCGCAAAACCGGAGAAATTTGACTCTCTGATGAGGTTAGC 118

QY 42 LysArgGluGlyIleAsnAlaGluGln\*\*\*ValIleLysIleThrAspGlnGlyTyrVal 61  
Db 119 AAGCGTGAAGGAATCAATGCTGAGCAATCGTCATCAAGATAACAGACCAAGGCTATGTC 178  
QY 62 ThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIleIle 81  
Db 179 ACTTCACATGGCGACCACTATCATTTATTAATGGAAGTTCCCTTATGACGCTATCATC 238  
QY 82 SerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleValAsn 101  
Db 239 AGTGAAGAATTACTCATGGAAGATCCAACTATAGCTAAAGATGAGGATATGTTAAT 298  
QY 102 GluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLysAsp 121  
Db 299 GAGGTCAAGGGTGGATATGTTATCAAGGTAGTAGGAAATATCTATGTTTACCTTAAGGAT 358  
QY 122 AlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnLysArg 141  
Db 359 GCTGCCCAACGCGGATAACGTCGTCACAAAGAGGAAATCAATCGACAAAACAGAGCAT 418  
QY 142 SerHisAsnHisAsnSer-----ArgAlaAspAsnAlaValAlaAlaArgAla 158  
Db 419 ACTCAACATCGTAGGTGAAGTCAAGAAACGATGGTCTGTGCTTGGCTTGGCAGCTGCG 478  
QY 159 GlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAsp 178  
Db 479 CAAGGAGCGCTACTACAGATGATGTTATCTTTAATGCTTCTGATATCATAGAGGAT 538  
QY 179 ThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGlu 198  
Db 539 ACTGCTGATGCTTATATCGTCTCATGAGGATCATTAACATACATTCCTTAAGATGAG 598  
QY 199 LeuSerAlaSerGluLeuAlaAlaGluAlaTyrTipAsnGly-----213  
Db 599 TTATCAGCTAGCGAGTTGGCTGCTCGAAGCCCTCTCTATCTGTCGAGGAAATCTGTCA 658  
QY 214 -----LysGlnGlySerArgProSerSerSerSerSerSerSerSerSerSer 227  
Db 659 AATTCAGGAACCTATCGCCGACAAATAGGATAACACTTCAAGAACAACTGGTACCT 718  
QY 228 AsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThrTyrHis 247  
Db 719 TCTGTGAAGCAATCCAGGAACCTACAAATACTAACACAGCAACCAACACACAACTAACAGT 778  
QY 248 Gln---AsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysProLeu 266  
Db 779 CAAGCAAGTCAAGTAATGACATGATGCTCTTGAACAGCTCTACAAACTGCCCTTTG 838  
QY 267 SerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArg 286  
Db 839 AGTCAACGACATGTAGATCTGATGCGCTTGTCTTGTATCCAGCAACAACTACAGTCA 898  
QY 287 ThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGluGln 306  
Db 899 ACAGCTAGAGGTGTGAGTGCACACAGGAGATCATCCACATTCATCCCTTACTCTCAA 958  
QY 307 MetSerGluLeuLysArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHis 326  
Db 959 ATGCTCGAATTTGAAGAACGAATCGCTCGTATTTATTTATTTCCCTTATCGTTCAAAACAT 1018  
QY 327 TrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSerPro 346  
Db 1019 TGGGTACCAAGATTCAAGGCCAGAACCAACAGTCCACACCGACTCCGGAACCTAGTCCA 1078  
QY 347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366  
Db 1079 GSCCCGCACTCGCAAAATCTTAA---ATAGACTCAAAAT-----TCTTCTTTG 1126  
QY 367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluAsnGlyValSer 386  
Db 1127 GTAGTACGCTGTCGAAAGTTGGGAGGATATGTATTCGAGAAAGGCGCATCTCT 1186



Db 1177 ATCAGTGAAGAGCTCCTCATGAAAGATCCGAAATATATCAGTTGAAGGATTCAGACATTGTC 1236  
Qy 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120  
Db 1237 AATGAAATCAAGGGTGTATGTGTCATTAAGGTAAACGGTAATAACTATGTTTACCTTAAG 1296  
Qy 121 AspAlaIleHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140  
Db 1297 GATGCAGCTCATGCGGATAATATTCGGAACAAAGAGAGATTAAACGTCAGAACGAGAA 1356  
Qy 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160  
Db 1357 CGCAGTCATATCACTCAAGAGCAGATATGCTGTGTCGAGCCGAGCCCAAGGA 1416  
Qy 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleLleGluAspThrGly 180  
Db 1417 CGTTATACACGGATGATGGGTATATCTTCAATGCACTGATATCATTTGAGACACGGGT 1476  
Qy 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200  
Db 1477 GATGCTTATATCGTTCCTCAGCGGACCAATACCATTTACATTCCTCAAGATGAGTTACA 1536  
Qy 201 AlaSerGluLeuAlaAlaGluAlaTyrTyrPheGlnGlyLysGlnGlySerArgProSer 220  
Db 1537 GCTAGCGAGTTAGCTGCTGCAGAGCCCTATGGAAATGGGAAGCAGGAGATCTCGTCTCT 1596  
Qy 221 SerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240  
Db 1597 TCAAGTCTAGTTATATGCAATCCAGCTCAACAGATTTGTACAGAAACCAATCTG 1656  
Qy 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260  
Db 1657 ACTGTCACCTCAACTATCATCAAAATCAAGGGGAAAAATTTCAAGGCTTTTACGTGAA 1716  
Qy 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280  
Db 1717 TTGTATGCTAAACCTTATCAGAACGCCATGTGAATCTGATGCTGATTTTTCGACCA 1776  
Qy 281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300  
Db 1777 GCGCAATCAAGTCGACCGCCAGAGGTAGCTGCTCCCTCATGTGTAACCATTTACAC 1836  
Qy 301 PheIleProTyrGluGlnMetSerGluLeuLysArgIleAlaArgIleIleProLeu 320  
Db 1837 TTTATCCCTTATGAACAAATGCTGTAATGGAAAAACGAATGCTCGTATTATTTCCCTT 1896  
Qy 321 ArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSer 340  
Db 1897 CGTTATCGTTCAAAACCATGCGTACCAAGATTCAGACCCAGAGAACCAAGTCCCAACCG 1956  
Qy 341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360  
Db 1957 ACTCCAGAACTAGTCCAAAGT-----CCGCAACCGAGCTCCAGCAAT 1998  
Qy 361 ProfileAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380  
Db 1999 CCAATTGATGAGAAATTTGGTCAAGAGAGCTGTCGAAAGATGAGCGATGGTTATGCTCTT 2058  
Qy 381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400  
Db 2059 GAGGAGAATGGAGTTCTCGTTATATCCAGCCCAAGGATCTTTTCAGCAGAAACAGCAGCA 2118  
Qy 401 GlyIleAspSerLysLeuAlaLysGlnLysSerLeuSerHisLysLeuGlyAlaLysLys 420  
Db 2119 GGCATTGATAGCAAACTGGCCCAAGCAGAAAGTTATCTCATAAGCTAGGAACTAAGAAA 2178  
Qy 421 ThrAspLeuProSerSerArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440  
Db 2179 ACTGACCTCCCTACTAGTATCGAGAAATTTCAATAGGCTTATGCTTACTACGAGA 2238  
Qy 441 IleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460  
Db 2239 ATTCAACCAAGATTTACTTGATAATAAAGGTCGACAAAGTTGATTTTGGGCTTTGGATAAC 2298

Qy 461 IeuLeuGluArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeu 480  
Db 2299 CTGTTGGACCACTCAAGGATGCTCAAGTGATAAAGTCAAGTTAGTGAAGATATCTTT 2358  
RESULT 10  
US-08-961-527-355  
; Sequence 355, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunach  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340PI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 355:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 973 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-527-355  
Alignment Scores:  
Pred. No.: 4,43e-122 Length: 973  
Score: 1377.00 Matches: 266  
Percent Similarity: 98.17% Conservative: 2  
Best Local Similarity: 97.44% Mismatches: 5  
Query Match: 34.26% Indels: 1  
DB: 4 Gaps: 0  
US-09-765-272A-66 (1-763) x US-08-961-527-355 (1-973)  
Qy 1 CysSerTyrGluLeuGlyValArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20  
Db 155 TGTTCTTATAGCTGGAGCTTACCAAGCTGTGTCAGGATGAGTAAGAGAGTCTAATCGATT 214  
Qy 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40  
Db 215 GCTTATATAGCTGGTGCATCAGCTGGTCCAAAGGCGCAAACTTGACACCATGAGTC 274  
Qy 41 SerLysArgGluGlyIleAsnAlaGluGln\*\*\*ValIleLysIleThrAspGlnGlyTyr 60  
Db 275 AGTAAGAGGGAGGGGATCAACCCGCAAAATTTGTTATCAAGATTACGGATCAAGGTTAT 334  
Qy 61 ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle 80  
Db 335 GTGACCTCTCATGGAGACCATTTATCTATTAATAAGGAGGTTCTTATGATGCATC 394





535 AlaTyrValThrProHisMetThrHisSerHisTrrpIleLysLysAspSerLeuSerGlu 554  
 3 GCCTATGTAACTCCACATATGACCATAGCACTGGATTAAAAAGATAGTTGCTGAA 62  
 555 AlaGluArgAlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProSerThr 574  
 63 GCTGAGAGAGCGGCACC-CAGGCTTATGCTTAAAGAGAAAGGTTTGACCCCTCCTTCGACA 121  
 575 AspHisGlnAspSerGlyAsnThrGluAlaLysGlyValAlaGluAlaIleTyrAsnArgVal 594  
 122 GACCATCAGGATTCAGGAATATCTAGGCAAAAGGACAGCATATCTACACCGCGTG 181  
 595 LysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGlu 614  
 182 AAAGCAGCTAAGAAGGTGCCACTTGATGCTTACCAATCTTCAATATCTAGTAA 241  
 615 ValLysAsnGlySerLeuIleProHisTyrAspHisTyrHisAsnIleLysPheGlu 634  
 242 GTCAAAACCGGTAGTTTAAATCATACCTCATATGACCATTACCATAACATCAAAATTGAG 301  
 635 TrpPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAla 654  
 302 TGGTTTGACAAAGCGCTTTATAGGCACCTTAAGGGGTATACTCTTGAGGATCTTTTGGCG 361  
 655 ThrValLysTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGly 674  
 362 ACTGTCAAGTACTATGTGCAACATCCAAACGACGTCGCGATTCAGATATGTTTGGT 421  
 675 AsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGlu 694  
 422 AACGCTAGCGACCATGTTCAAGAAACAAAAATGGTCAAGCTCATCAACATCAACCGAA 481  
 695 LysProSerGluGluLysProGlnThrGluLysProGluGluLysProArgGluGlu 714  
 482 AAACCAAGCGAGAGAACTCAGACAGAAACCTGAGAGAAACCTGAGAGAAACCTGAGAGAG 541  
 715 LysProGlnSerGluLysProGluSerProLysProThrGluGluProGluGluSerPro 734  
 542 AAACCGCAAGCGAGAAACCAAGCTCTCAAAACCAACCAAGAGAACCAAGAAATCAACA 601  
 735 GluGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArgGluAla 754  
 602 GAGGAATCAGAAAGAACCTCAGGTCGAGACTGAAAGGTTGAAGAAACCTGAGAGAGGCT 661  
 755 GluAspLeuLeuGlyLysIleGlnAsp 763  
 662 GAAGATTACTTTGGAAAAATCCAGGAT 688

## RESULT 13

JS-09-468-656A-7

Sequence 7, Application US/09468656A

Patent No. 6582706

GENERAL INFORMATION:

APPLICANT: Johnson, Leslie S.

APPLICANT: Adamou, John E.

TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus

TITLE OF INVENTION: Pneumoniae Group A and B Having Selected Structural

TITLE OF INVENTION: Motifs

FILE REFERENCE: 469201-444

CURRENT APPLICATION NUMBER: US/09/468,656A

PRIOR FILING DATE: 1999-12-02

PRIOR FILING DATE: 60/113,048

PRIOR FILING DATE: 1998-12-21

NUMBER OF SEQ ID NOS: 14

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 7

LENGTH: 1455

TYPE: DNA

ORGANISM: Streptococcus pneumoniae

JS-09-468-656A-7

Alignment Scores:

Pred. No.:	8.48e-102	Length:	1455
------------	-----------	---------	------

Score: 1167.50 Matches: 254  
 Percent Similarity: 59.50% Conservative: 78  
 Best Local Similarity: 45.52% Mismatches: 123  
 Query Match: 29.05% Indels: 103  
 DB: 4 Gaps: 14

US-09-765-272A-66 (1-763) x US-09-468-656A-7 (1-1455)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20  
 DB 61 TGTGCTTATGCACTAAACACGATCGTTGCG---CAGGAAAAATAAGCAATATATCGTGTG 117  
 QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40  
 DB 118 TCTTATGTGATGTCAGCCAGTCAGTCAAGTCAAGAAAGTGAACCTTGACACGACGAGTT 177  
 QY 41 SerLysArgGluGlyIleAsnAlaGluGln\*\*ValIleLysIleThrAspGlnGlyTyr 60  
 DB 178 AGCCAGAAAGAAAGAAATTCAGGCTGAGCAAAATTGTAATCAAAATTTACATCAGGCGCTAT 237  
 QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80  
 DB 238 GTAAACGTCAACGCGTGACCTATCATTAATATGGAAGAGTTTCTTATGATGCGCTTC 297  
 QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100  
 DB 298 TTTAGTGAAGAACTCTTGATCAAGATCCAACTATCAACTTAAAGACCTGATATTGTGTC 357  
 QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120  
 DB 358 AATCAAGTCAAGGCTGGTTATATCATCAAGTTCGATCGAAATATATTGCTTACCTGAAA 417  
 QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140  
 DB 418 GATCAGCTCATGCTGATATATGTTGCAACTAAGATGAATCAATCGTCAAAACAAGAA 477  
 QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaArgAlaGlnGly 160  
 DB 478 CATGCTCAAGAT---AATGAAAGGTTAACTCTATGTTGCTGTAGCAAGGCTCTCAGGGA 534  
 QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180  
 DB 535 CGATATACGCAAAATGATGTTATGCTTTTAAATCCAGCTGATATATTATCGAAGATACGGT 594  
 QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200  
 DB 595 AATGCTTATATCGTTCTCTCATGAGGTCATCATCACTACATTCCTCCAAAAGCGATTATCT 654  
 QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArg 218  
 DB 655 GCTAGTGAATTAGCAGCAGCTTAAAGCACATCTGGCTGGAAAAAATATATCAACCGAGTCA 714  
 QY 219 ProSerSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238  
 DB 715 TTTAGCTATTCTTCAACAGCTAGTAGCAATCAACGCAA-----TCTGTACAAAA 765  
 QY 239 AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeu 258  
 DB 766 GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAAATCTCCAGAGTCTTTTG 816  
 QY 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePhe 278  
 DB 817 AAGGAACCTCATGATTCACTCAGCCCAAGTTACAGTGAATCAGATGGCTGGTCTTT 876  
 QY 279 AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis 298  
 DB 877 GACCTGCTAAGATTATCAGTCGTACACCAATGGAGTTGCGATTCGCCATGGCGACCAT 936  
 QY 299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIle 318  
 DB 937 TACCACCTTTATCTTACAGCAAGCTTCTGCTTAGNAGAAAGATGCGCAGATGGTG 996  
 QY 319 ProLeuArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerPro 338

```
Db 997 CCTATCAGTGAACCTGGTCTACAGCTTTCTACAAATGCARAAACCTAAAT----- 1044
Qy 339 GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro 358
Db 1045 -----GAAGTAGTGTCTAGTCTAGGCGAGTCTTCA 1074
Qy 359 SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378
Db 1075 AGCAATCCTCTCTTTACAGCAAGTAGAGG---CTCTCTTCAGCATCTGATGGTTAT 1131
Qy 379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLysSerAlaGluThr 398
Db 1132 ATTTTAAAT-----CCAAAAGATATCTGTTGAAGAAACG 1164
Qy 399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLysSerHisLysLeuGlyAla 418
Db 1165 GCTACA-----GCTTATATTGTAGACATGCTGATCATTTCCATTAC----- 1206
Qy 419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438
Db 1207 -----ATTCCAAATCAAAT----- 1221
Qy 439 AlaArgIleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458
Db 1222 ---CAAATTTGGCAACCGACTCTTCCAAACAATAGCTTAGCAACACCTTCTCCATCTCT 1278
Qy 459 AspAsnLeuLeuGluArgLysLeuVal***SerAspLysValLysLeuVal***Asp 478
Db 1278 ----- 1278
Qy 479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498
Db 1279 -----CAATC-----AATCCAGGA 1293
Qy 499 IleThrTyrThrAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGlu 518
Db 1294 ACTTCACATGAGAAACATGAA-----GAA 1317
Qy 519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
Db 1318 GATGGATACGGATTGTATGCTTAATGCTTAATGCTGAAGATGAATCAGGTTTTCATG 1377
Qy 539 ProHisMetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGlu 556
Db 1378 AGTCACGAGACCAATCAATATTTCTTCAGAAAGGACTTGACAGAAAGCA 1431
```

## RESULT 14

```
US-08-961-083-181
; Sequence 181, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 181:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1342 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-083-181

Alignment Scores: 2,26e-99 Length: 1342
Pred. No.: 1141.50 Matches: 250
Score: 59.41% Conservative: 75
Percent Similarity: 45.70% Mismatches: 119
Best Local Similarity: 28.40% Indels: 103
Query Match: 3 Gaps: 14
DB:

US-09-765-272A-66 (1-763) x US-08-961-083-181 (1-1342)

Qy 5 LeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSerTyrIleAsp 24
Db 2 CTAACCCAGCATCGTTCG---CAGGAAATATAGGACAATAATCGTGTCTCTTATGTGGAT 58
Qy 25 GlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluValSerLysArgGlu 44
Db 59 GGCAGCCATCGAGTCAGAAAGTGAAGAACTTCACACCAGCCAGGTTAGCCAGAAAGAA 118
Qy 45 GlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyrValThrSerHis 64
Db 119 GCAATTCAGGCTGAGCAAAATTTGTAATCAAAATTTACAGATCAGGCTATGTAAAGTCAC 178
Qy 65 GlyAspHisTyrHisTyrTrpAsnGlyLysValProTyrAspAlaIleIleSerGluGlu 84
Db 179 GGTGACCATCTATCTACTATATATATGGAAGTTCTTATGATGCCCTCTTTAGTGAAGAA 238
Qy 85 LeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleValAsnGluIleLys 104
Db 239 CTCCTGATGAAGATCCAACTATCACTTAAGAGCTGATATTGTCAATGAGTCAAG 298
Qy 105 GlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLysAspAlaHis 124
Db 299 GGTGGTTATATCATCAAGGTCGATGCAAAATATTTATGTCTTACCTGAAAGATGCAAGTCAT 358
Qy 125 AlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGluArgSerHisAsn 144
Db 359 GCTGATAATGTTGCAACTAAAGATGAATCAATCGTCAAAACAAGAACATGTCAAAGAT 418
Qy 145 HisAsnSerArgAlaAspAsnAlaValAlaAlaArgAlaGlnGlyArgTyrThrThr 164
Db 419 ---AATGAGAGGTTAACTCTAATGTTGCTAGCAAGGTTCTCAGGACCATATACACA 475
Qy 165 AspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIle 184
Db 476 AATGATGGTTATGTTCTTTAATCCAGCTGATATTTATCGAGATACGGGTAAATGCTTATATC 535
Qy 185 ValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeu 204
Db 536 GTTCTCTCATGGAGGTCACATCTACTACTATCCAAAGGCGATTATCTGCTAGTGAATTA 595
Qy 205 AlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArgProSerSer 222
Db 596 GCAGCAGCTAAAGCACATCTGCTGGAAAAAATATGCAACCGAGTCAGTTAAGCTATTCT 655
Qy 223 SerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrVal 242
Db 656 TCAACAGCTAGTGACAATAACACGCA-----TCTGTACCAAAAGGATCACTACGCTAGC 706
```





145 HisAsnSerArgAlaAspAsnAlaValAlaAlaAargAlaGlnGlyArgTyrThr 164  
149 ---AATGAGAAGGTTAACTTAATGTTGCTGTAGCAAGGCTCTCAGGACGATATACGACA 475  
165 AspAspGlyTyrIlePheAsnAlaSerAspIleGluAspThrGlyAspAlaTyrIle 184  
170 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
185 ValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeu 204  
190 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
205 AlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArgProSerSer 222  
210 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
223 SerSerTyrAsnAlaAlaProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrVal 242  
247 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
256 TCAACACAGCTAGTACCAATAACACGCA-----TCTGTAGCAAAAGGATCAACTAGC 706  
263 ThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyr 262  
267 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
270 ARGCCAGCA-----AATAATCTGAAATCTCCAGAGCTCTTTGAAGGAACCTCTAT 757  
283 AlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGln 282  
287 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
298 GATTCACCTAGCCCAACCTTACAGTGAATCAGATGCGCTGCTGTTGACCTGCTTAAG 817  
303 IleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIle 302  
307 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
318 ATTATCAGTCTGACCAAAATGAGTTCGATTCGCGATCGCGACCATTAACCACTTATT 877  
323 ProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyr 322  
327 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
338 CCTTACAGCAAGCTTTCTGCTTGTAGAAAGAAAGATGCGAGATGGTGGCTTATCAGTGA 937  
343 ArgSerAsnHisTrpValProAspSerArgProGluGlnProSerProGlnSerThrPro 342  
347 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
358 ACTGGTCTACAGTTCTACAAATGCAAAACCTAAT----- 973  
363 GluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIle 362  
367 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
374 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
383 AspGluLysLeuValLysGluAlaValAlaArgLysValGlyAspGlyTyrValPheGluGlu 382  
387 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
396 TCTTTAACGCAAGTAAGGAG---CTCTCTCAGCATCTGATGTTATATTTTAAAT--- 1069  
403 AsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIle 402  
407 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
410 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
423 AspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAsp 422  
427 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
430 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
443 LeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHis 442  
447 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
456 ATTCCAAATCAAAAT-----CAAATTTGGG 1159  
463 GluAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeu 462  
467 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
470 CAACCGACTCTTCCAAACAATAGTCTAGCAACACTTCTCCATCTCTT----- 1207  
483 GluArgLeuLysValVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeuAlaPhe 482  
487 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
490 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
503 LeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThr 502  
507 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
510 ---GCTTATGCTGCTCTTATCAGCTGATATATCAGGATGCGGTAATGCTTATATC 535  
523 AspAspGluLeuGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIle 522

1235 AAACATGAA-----GAAGATCGATACGGA 1258  
523 PheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThr 542  
1259 TTTGATGCTAATCGTATTATTCGCTGAAGATGAATCAGGTTTGTGATGAGTACGAGAC 1318  
543 HisSerHisTrpIleLysLys 549  
1319 CACAATCATTTATTTCTTCAAG 1339  
RESULT 16  
US-08-743-637B-34/c  
; Sequence 34, Application US/08743637B  
; Patent No. 5994066  
; GENERAL INFORMATION:  
; APPLICANT: BERGERON, Michel G.  
; APPLICANT: PICARD, Francois J.  
; APPLICANT: OUELLETTE, Marc  
; APPLICANT: ROY, Paul H.  
; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA  
; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND  
; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED  
; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...  
; NUMBER OF SEQUENCES: 273  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: QUARLES & BRADY  
; STREET: 411 EAST WISCONSIN AVENUE  
; CITY: MILWAUKEE  
; STATE: WISCONSIN  
; COUNTRY: USA  
; ZIP: 53202-4497  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/743,637B  
; FILING DATE: 04-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/526,840  
; FILING DATE: 11-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BAKER, Jean C.  
; REGISTRATION NUMBER: 35,433  
; REFERENCE/DOCKET NUMBER: 850586.90012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (414) 277-5000  
; TELEFAX: (414) 277-5591  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 841 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pneumoniae  
US-08-743-637B-34  
Alignment Scores:  
Pred. No.: 1,368-39 Length: 841  
Score: 512.00 Matches: 104  
Percent Similarity: 93.10% Conservative: 4  
Best Local Similarity: 89.66% Mismatches: 2  
Query Match: 12.74% Indels: 7  
DB: 2 Gaps: 2  
US-09-765-272A-66 (1-763) x US-08-743-637B-34 (1-841)  
QY 653 LeuAlaThrValLysTyrThrValGluHisProAsnGluArgProHisSerAspAsnGly 672

Db 839 TTGGCGACTGTCAAGTACTATGTCGAACATCCAAACGACGTCG-CATTGAGATAATGGT 781  
Qy 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692  
Db 780 TTGTGTAACGCTAGCGACCATGTTCAAGAAACAAATAATGGTCAAGCTGATACCAATCAA 721  
Qy 693 ThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluGluThrProArg 712  
Db 720 ACGGAAAAACCAACGAGGAGAACTTCAGACAGAAAACCTCAGGAGAAACCCCTCGA 661  
Qy 713 GluGluLysProGlnSerGluLysProGluSerProLysProThrGluGlu----- 729  
Db 660 GAAGAGAAACCGCAACGAGGAGAAACAGAGTCTCCAAAACCAACAGAGGAAACCAAGAA 601  
Qy 730 -----ProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLysVal 747  
Db 600 GAATCAACGAGAAATCACAGAGGAATCAGAGAACCTCAGGTCGAGACTGAAAGGTTA 541  
Qy 748 GluGluLysLeuArgGluAlaGluAspLeuLysGlyLysIleGlnAsp 763  
Db 540 ---AGAAACTGAGAGGCTGAGATTACTTGGAAAAATCCAGAT 496  
RESULT 17  
US-08-526-840B-34/c  
; Sequence 34, Application US/08526840B  
; Patent No. 6001564  
; GENERAL INFORMATION:  
; APPLICANT: BERGERON, Michel G.  
; APPLICANT: OUELLETTE, Marc  
; TITLE OF INVENTION: SPECIFIC AND UNIVERSAL PROBES AND  
; TITLE OF INVENTION: AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND IDENTIFY  
; TITLE OF INVENTION: COMMON BACTERIAL PATHOGENS AND ANTIBIOTIC RESISTANCE GENES  
; TITLE OF INVENTION: FROM CLINICAL SPECIMENS FOR ROUTINE DIAGNOSIS IN ...  
; NUMBER OF SEQUENCES: 177  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: QUARLES & BRADY  
; STREET: 411 East Wisconsin Avenue  
; CITY: Milwaukee  
; STATE: Wisconsin  
; COUNTRY: USA  
; ZIP: 53202-4497  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/526,840B  
; FILING DATE: 11-SEP-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/304,732  
; FILING DATE: 12-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BAKER, Jean C.  
; REGISTRATION NUMBER: 35,433  
; REFERENCE/DOCKET NUMBER: 850586.90012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (414) 277-5000  
; TELEFAX: (414) 277-5591  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 841 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pneumoniae  
US-08-526-840B-34

Alignment Scores:  
Pred. No.: 1.36e-39 Length: 841  
Score: 512.00 Matches: 104  
Percent Similarity: 93.10% Conservative: 4  
Best Local Similarity: 89.66% Mismatches: 2  
Query Match: 12.74% Indels: 7  
DB: 3 Gaps: 2  
US-09-765-272A-66 (1-763) x US-08-526-840B-34 (1-841)  
Qy 653 LeuAlaThrValLysTyrValGluHisProAsnGluArgProHisSerAspAsnGly 672  
Db 839 TTGGCGACTGTCAAGTACTATGTCGAACATCCAAACGACGTCG-CATTGAGATAATGGT 781  
Qy 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692  
Db 780 TTGTGTAACGCTAGCGACCATGTTCAAGAAACAAATAATGGTCAAGCTGATACCAATCAA 721  
Qy 693 ThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluGluThrProArg 712  
Db 720 ACGGAAAAACCAACGAGGAGAAACCTCAGACAGAAAACCTCAGGAGAAACCCCTCGA 661  
Qy 713 GluGluLysProGlnSerGluLysProGluSerProLysProThrGluGlu----- 729  
Db 660 GAAGAGAAACCGCAACGAGGAGAAACAGAGTCTCCAAAACCAACAGAGGAAACCAAGAA 601  
Qy 730 -----ProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLysVal 747  
Db 600 GAATCAACGAGAAATCACAGAGGAATCAGAGAACCTCAGGTCGAGACTGAAAGGTTA 541  
Qy 748 GluGluLysLeuArgGluAlaGluAspLeuLysGlyLysIleGlnAsp 763  
Db 540 ---AGAAACTGAGAGGCTGAGATTACTTGGAAAAATCCAGAT 496  
RESULT 18  
5231168-1  
; Patent No. 5231168  
; APPLICANT: DIEGIEL, MORTEN; BORRE, MARTIN; JEPSEN, SOREN;  
; VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.  
; TITLE OF INVENTION: MALARIA ANTIGEN  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/409,658  
; FILING DATE: 18-SEP-1989  
; SEQ ID NO: 1:  
; LENGTH: 3095  
5231168-1  
Alignment Scores:  
Pred. No.: 1.05e-06 Length: 3095  
Score: 177.00 Matches: 136  
Percent Similarity: 36.19% Conservative: 109  
Best Local Similarity: 20.09% Mismatches: 234  
Query Match: 4.40% Indels: 198  
DB: 6 Gaps: 34  
US-09-765-272A-66 (1-763) x 5231168-1 (1-3095)  
Qy 177 GluAspThrGlyAspAlaTyrIleValProHisGly-----AspHisTyr 191  
Db 103 GAAGAATCAGGTGAAAGTAAATAGTTGATAATGATGAAGTGGTTTGAAGAAAGCTCAT 162  
Qy 192 HistyrlleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrp 211  
Db 163 CATGAAAAATTTTCATCTGAAAGTAACTCTGAATTAATAATGAAATGAAATTTGTTGAA 222  
Qy 212 AsnGlyLysGlnGlySerArgProSerSer-----SerSerSerTyrAsn 226  
Db 223 TCTGACAAAAGTGTAACTGAACTCTGAAATGAAAGAGTGTGTATCTGAAGAAAGCAAC 282  
Qy 227 AlaAsnProAla-----GlnProArgLeuSerGluAsnHisAsnLeuThrValThr 243  
Db 283 CCTGAAACCGCTGAAAAATGAAAGAGTAGTAGAGAGAGCTCATCAGGAGAAAGATTGTA 342

244	QY	ProThrTyrHisGlnAsnGlnGlyIleuAsnHisSerSerLeuLeuArgGluLeuTyrAla	263
343	DB	CCT-----GAACAAATGATGAAGAATCAGTGCAGATGGATTAGTTGATAATGAAGAA	396
264	QY	LysProLeuSerGluArgHisValGluSerAspGlyLeuLeuPheAspProAlaGlnIle	293
397	DB	GGTGATTTTGAAGAACCTAATCATGAAGAA-----TTTGAACTGATCAAAAT	444
284	QY	ThrSer-----ArgThrAlaArgGlyValAlaValProHis	295
445	DB	GACTCTGAATTAAGTGAAATGAATAGTTGAATCAGAAAAAAGTGATCTGAACCA	501
296	QY	GlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGluIysArgIleAla	315
502	DB	GCTGAACATGTGAAATTTGATCAGAAAAAAGTGATCTGAAACAGCTGAACAGTA	558
316	QY	ArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGln	335
559	DB	GAAATTTGTATCTGAAAAAGTACATCCGAA-----CCAGCTGAACATGTAGAAAGT	609
336	QY	ProSerProGlnSerThrProGluProSerProSerProGlnProAlaProAsnProGln	355
610	DB	GTATCTGAACAAAGTAATAACGAAACCTCCGAAAG-----AAAGATGGA	654
356	QY	ProAlaProSerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGly	375
655	DB	CCAGTTTCCTTCAAAACCATTTGAAGAAATGAAAAAGTGGATGTTCAACCTAAATTTGA	714
376	QY	AspGlyTyrValPheGluGluAsnGlyValSer-----	386
715	DB	GACCTTCAAAATAATTGAACCTAAATTTGTGTGACTCACAACCAATCCACAAGAACCA	774
387	QY	-----ArgTyrIleProAlaLysAspLeuSerAlaGluThrAla	399
775	DB	GAACCATCAATTTGTCAAAATTTGAAAAAGTTCCTTCTGAAGAA-----ATAAACAT	825
400	QY	AlaGlyIleAspSerLysLeuAlaLysGlnGluSerHisLysLeuGlyAlaLys	419
826	DB	GCAAGTGTGATCCTCGAAGTAAAGAAAAAGAAATGCTATCTGAAGTTGTTGAAGAAA	885
420	QY	LysThr-----AspLeuProSerSerAspArgGluPheTyrAsnLys	433
886	DB	CAAATTTCCAGAAATCAGTTGGAAGAAATTCAGTAAATCAGGATCAATTTGAAGAT	942
434	QY	AlaTyrAspLeuAlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnVal	453
943	DB	-----GTTCACTCGAACCAATTAGAT-----TTA	966
454	QY	AspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysVal	473
967	DB	GATCATAAACAGTTGTATCCAGAAATAGTAGAAGTTGAAGAAATTCCTTCAGACATACAT	1026
474	QY	LysLeuVal***AspIleLeuAlaPheLeuAlaProIleArgHisProGluArgLeu	492
1027	DB	GAA-----AATCAAGTGGCTCATCCAGAAATTTGTTGAA	1059
493	QY	-----GlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIleGln	507
1060	DB	ATTGAGGAAGTTTCTCTGAACCAAAAT-----CAAAATAACGAATTTCAA	1104
508	QY	ValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAspProArgAsp	527
1105	DB	-----GAA	1107
528	QY	IleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTrpIle	547
1108	DB	ATTAATGAAGATGATAAAAGTGCA-----CATATTCAGCAT-----	1143
548	QY	LysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAlaLysGluLys	567
1144	DB	-----GAAATAGTAGAAGTAGAGAAATACTTCCGAAGATGATAAAATGA	1194

568 GlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGlyAla 597  
1195 -----GTTGAACATGAA-----  
588 GluAlaIleTyrAsnArgValLysAlaAlaLysValProLeuAspArgMetProTyr 607  
1207 -----ATAGTAGAAGTTGAAGAATTTCTACCAAGAAGATAAAATGAAGAAA 1251  
608 AsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleIleProHisTyrAspHis 627  
1252 GGTCAACATCAAAATAGTAGAGTTGAAGAA-----ATTCTACCAGAAGATGATAAA 1302  
628 TyrHisAsnIleLysPheGlu-----TTPPhAspGluGlyLeuTyrGluAlaProLys 645  
1303 AATGAAAGAGTTGAACATCAAAATAGTAGAAGTTGAAGAAATTTCTACCAAGAAGATAAAAT 1362  
646 GlyTyrThrLeuGluAspLeuAlaThrValLysTyrThrValGluHisProAsnGlu 665  
1363 GAAAAAGGTCAACATGAATAGTAGAGTTGAAGAAATTTCTACCAAGAAGATAAAATGAA 1422  
666 Arg-----ProHisSerAspAsnGlyPhe 673  
1423 AAGTTCAACATGAATAGTAGAAGTTGAAGAAATTTCTACCAAGAAGATAAAATGAAGAAA 1482  
674 Gly-----AsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAla 688  
1483 GGTCAACATCAAAATAGTAGAGTTGAAGAAATTTCTACCAAGAAGATAAAATGAAGAGTT 1542  
689 Asp-----ThrAsnGlnThrGluLysProSer 697  
1543 CAACATGAATAGTAGAAGTTGAAGAAATTTCTACCAAGAAGATAAAATGAAGAGTCAA 1602  
698 GluGluLysProGlnThrGluLys-----ProGluGluGluThrProArgGluGluLys 715  
1503 CATGAATAGTAGAGTTGAAGAAATTTCTACCAAGAAGAGAT-----AAAAATGAAGAAA 1656  
716 ProGlnSerGluLysProGluSerProLysProThrGluGlu-----ProGluGluSer 733  
1657 GGTCAACATCAAAATAGTAGAG-----GTTGAAGAAATTTCTACCAAGAAGATAAA 1704  
734 ProGlu-----GluSerGluGlu-----ProGlnValGluThr 744  
1705 AATGAAAGAGTTCAACATGAATAGTAGAGTTGAAGAAATTTCTACCAAGAAGATAAAAT 1764  
745 GluLysValGluGluLysLeuArgGluAlaGluAspLeuGlyLysIle 761  
1765 GAAAAAGTTCAACATGAATAGTAGAGTTGAAGAAATTTCTCCAGAAATTTCTCCAGAAAT 1815

RESULT 19  
US-08-257-073-4  
; Sequence 4, Application US/08257073  
; Patent No. 5766597  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: de Taisne, Charles  
; APPLICANT: Tine, John A.  
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
; NUMBER OF SEQUENCES: 143  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue, 25th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/257,073  
; FILING DATE: 09-JUN-1994









512 ACSCGCTCTCGCATGGCATGATTGACCGAGTCCCGCTTCTGCATCTCCGACGAGAG 571  
416 LeuGlyAlaLys---LysThrAspLeuProSer-----SerAspArgGluPheTyr 431  
572 GTCGGCTCGGACGTTTCCATCGAGTCCGTTCCGAATGCGGTTTCCCTCGCGGAGTCCGC 631  
432 AsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeuAsnLysGly--- 450  
632 GAGTGTGTGAACCGCTCTCAGGAGAGAGCTACTCGGACTGCCCGCATCATAGGAATC 691  
451 ArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLysAspVal\*\*\*Ser 470  
692 CGTCAGTGTCTGACGAATCTGCTCTGCTCGCTTCGGCGAG----- 730  
471 AspLysValLysLeuVal\*\*\*AspIleLeuAlaPheLeuAlaProIleArgHisProGlu 490  
731 -----TGTCACCTCGCGGGAACCCAG 754  
491 -----ArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAsp 503  
755 CAAGGCTCGGTATCCGCAAGAGACGTGCATCGCAACAGCTGCTGCGCTGTGTCCAG 814  
504 AspGluIle-----GlnValAlaLysLeuAlaGlyLys---TyrThr 516  
815 GCGGAGGTCTGCGCGGATGTCACCCAGAGATTGAGGAGGAAGAGCAACATTTCCCC 874  
517 ThrGluAspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAsp----- 534  
875 CCGAGAGAGCGAGGCTTTCCTCCATATGAAGAGGGTCTGTTGAGGTGAGCTGTT 934  
534 ----- 534  
935 CTTCCCGAGGAGGATPCCCTGAAGAGAAACATCTCTGAGGAGGAAATCCCTGAAGGA 994  
535 AlaTyrValThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGlu 554  
995 GAACATATTTCTGAAGAGCTCCCAAGAGGCGAGCATGTTCTCTGAGGAGAAATCCCTGAA 1054  
555 AlaGluArgAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProSerThr 574  
1055 GGAGACATGTTCTGAGAGGAATC-----CCTGAAGGA 1090  
575 AspHisGlnAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgVal 594  
1091 GAGCATGTTCTGAAGAGTTCCTCCAGAA-----GGCGAACATGTTCTGAGGAGGAAATC 1144  
595 LysAlaAlaLysLysValProLysAspArgMetProTyrAsnLeuGlnTyrThrValGlu 614  
1145 CCGTGAAGGAGAACATGTTCTCTGAAGAGGAAATCCCTGAAGGAGAACATGTTCTCTGAAGAG 1204  
615 ValLysAsnGlySerLeuIle-----IleProHisTyrAspHisTyrHisAsnIle 631  
1205 TTTCCCTGAAGGAGAACATATTTCTGAGGAGCTCCTGAGGAGGAGCATATTCAT----- 1258  
632 LysPheGluTrpPheAspGluGlyLeuTyr-----GluAlaProLysGly---Tyr 647  
1259 -----GAGAGGTTCCCTGAAGGAGAGCATATTCATGAGGAGCTCCTGAGGAGGAGCAT 1312  
648 ThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGluArgPro 667  
1313 GTTCTCTGAAGAGGAGATCCCTGAAGGAGAACATATTCCTGAGGAGTTCCTCTGAGGAGGAG 1372  
668 HisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGln 687  
1373 CATGTTCTGAGGAGGAAATCCCTGAAGGAGAACATATTCCTGAGGAGGAGTTCCTCTGAA 1432  
688 AlaAspThrAsnGlnThrGluLys-----ProSerGluLys 700  
1433 GGAGAGCATGTTCTGAGGAGGAGATCCCTGAAGGAGGAGCATGTTCTGAGGAGGAGCTC 1492  
701 ProGlnThrGluLys---ProGluGluGluThrProArg----- 712  
1493 CTTGGAGGAGAACTTATTCTCTGAGGAGGAGATCCCTGAAGGAGGAGCATGTTCTCTGAAGAG 1552

QY 713 -----GluGluLysProGlnSerGluLysProGluSerProLysProThrGlu 728  
Db 1553 CTCCTGAAGCGAGCATGTTCTGAGGAGGAGATCCCTGAAGGAGGAGCATGTTCTCTGAA 1612  
QY 729 Glu-----ProGluGluSerProGluGluSerGluGlu 739  
Db 1613 GAGGAATTCCTGAAGCGGAGCATGTTCTGAGGAGGAGACCTCTGAAGGAGAACATGCT 1672  
QY 740 ProGlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAsp 756  
Db 1673 CCAGGAAGAGACTCTCTGCACCTGAGGAGACCGGAAAGAGGAGGAAGAA 1723

## RESULT 22

US-08-242-932-1  
; Sequence 1, Application US/08242932  
; Patent No. 5595740  
; GENERAL INFORMATION:  
; APPLICANT: Brady, L. Jeanine  
; TITLE OF INVENTION: Cloning of No. 5595740-Iga Fc Binding Forms of  
; TITLE OF INVENTION: the Group B Streptococcal Beta Antigens  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/242,932  
; FILING DATE: 16-MAY-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: UFI42  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4200 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-242-932-1

## Alignment Scores:

Pred. No.:	3,73e-05	Length:	4200
Scores:	163.00	Matches:	166
Percent Similarity:	35.26%	Conservative:	127
Best Local Similarity:	19.98%	Mismatches:	316
Query Match:	4.06%	Indels:	222
DB:	1	Gaps:	44

US-09-765-272A-66 (1-763) x US-08-242-932-1 (1-4200)

QY 8 HisGlnAlaGlyGlnValLysLys-----GluSerAsnArgValSerTyrIleAspGly 25  
Db 917 CATCAAAAAGTCAAGTTGAAAAAATGCGAGAGCAAAAGGGAATCACAATGAAGATAAA 976  
QY 26 Asp-----GlnAlaGlyGlnLysAlaGluAsnLeu 35  
Db 977 GATTCTATGCTGAAAAAATCGAAGATATTCGTAACAAGCTCAACAGCAGATAAAAA 1036  
QY 36 ThrProAspGluValSerLysArgGlu----- 44

1037 GAAGATGCGGAAGTAAAGGTTTCGTGAAGAACTAGGTAAACTCTTTAGTTCAACTAAAGCT 1096  
 45 GlyIleAsnAlaGlu-----Gln\*\*ValIleIysIleThrAspGlnGlyTyrValThr 62  
 1097 GGTCTGGATCAAGAAATTCAGAGCATGTGAAGAAAGAAACAGTAGTGAGGAAATATCT 1156  
 63 SerHisGlyAspHisTyrHisTyrTyrAsn-----GlyLys 74  
 1157 CAGAAAGTTGATGAA--CACTATGCTAATAGCCTTCAGAACTTCGCTCAAAAATCTCTT 1213  
 75 ValProTyrAspAlaIleIleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeu 94  
 1214 GAAGAACTAGATAGGCAACTACCAATGAACAAGCTACACAAGTTTAAAAATCAATCTCTTA 1273  
 95 LysAspSerAspIleValAsnGlnIleLysGlyGlyTyrValIleLysValAsnGlyLys 114  
 1274 GAAACGCTCAAAAGCTCAAAAGATAACAA--CCTCTTATCAAAAGAACGAATGTGAAA 1330  
 115 TyrTyrValTyrLeuLysAspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIle 134  
 1331 TTGTATAGGCTATG-----AGTGAGAGCTTG 1357  
 135 LysArgGlnLysGlnGluArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAla 154  
 1358 GAGCAGGTTGAGAGGAATTAATAACATAATTCGGAAGCTAAATTTAGAACGATTTGGTTGG 1417  
 155 AlaAlaArgAlaGlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAsp 174  
 1418 AAATCTAAA-----GAA 1429  
 175 IleIleGluAspThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIle 194  
 1430 ATCGTAAGAA-----TACGAGGAAAA 1453  
 195 ProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLys 214  
 1454 CTTAATCAATCTAAAATCTTCAGATTTAAAGCACTAGAGAGGAGGAGCTCATTCGAG 1513  
 215 -----GlnGlySerArgProSerSerSerSerSerTyrAsnAlaAsnPro 229  
 1514 TTGAACAAGTTGTGGAGATTTTGAAGAAAAATTTAAACCGTCAGAGCAAGTGACACCA 1573  
 230 -----AlaGlnProArgLeuSerGluAsnHisAsn-----Leu 240  
 1574 AAAAACGTTCAACGAGATTAGCTGTCTAATGAAAAATTAATCAACAAAGATTGAGTTA 1633  
 241 ThrValThrProThrTyrHisGlnAsn-----GlnGlyGluAsnIleSerSer 256  
 1634 ACAGTTTCACCA-----CAGAAATATCATGTATATGAAGTGAAGCGTGAATTT 1684  
 257 LeuLeuArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeu 276  
 1685 ACAGTCACA-----GCTAAAAGTGTATCGAAGACGACGCTTGCACTTCAGTGATCTT 1735  
 277 Ile-----PheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValPro 294  
 1736 TTACAAATATTAATCCGCTCT--GTATCAGATAGAAATTAGTACAAATTTATAGACTAAC 1792  
 295 HisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGlyLysArgIle 314  
 1793 ACGGATATCATAG-----ATTGCCGAATCATCATCTCAAGNATTTGAAGCTAAATGAA 1846  
 315 AlaArgIleIleProLeuArgTyrArgSerAsn-----His 326  
 1847 AGTCAACAGTCACTCTAAAAGCTAAAGATGATTCGCAATCTAGTTGAAAAACATTC 1906  
 327 TrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSerPro 346  
 1907 ACTATTACAGTCAAAAGAGAGGAGAGAACAAAGTTCTCTAAAACACCAAGCAGAGAGAT 1966  
 347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366

Db 1967 TCTAAACCGAAGAAAGGTTCTCTCAAGAACCAAAATCAAAAT-----GACAAGAAATCAA 2020  
 QY 367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386  
 Db 2021 TTCAAGAGTTGATTAATCAATCAGCTCAACAGAACTGGAAGAGTTAGAAAAAGCAATAAAA 2080  
 QY 387 ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406  
 Db 2081 GAATTAATGGAGCAACCCAGAGATTCATCCATCCAGAGTATGGTATTCAAAAATCTATT 2140  
 QY 407 -----AlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuPro 424  
 Db 2141 TGGAGGTCACAAAAGAGCGCTATCCAGAAAGCCATAACAGCTTTTAAAGAGATTATGGT 2200  
 QY 425 SerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAsp 444  
 Db 2201 GATTCATCTTCAAAATACTACACAGACACTAT----- 2233  
 QY 445 LeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArg 464  
 Db 2234 -----TTTAAACAATATAATCT--GATTTT-----ATGAATTATCACTTCATGCA 2278  
 QY 465 LeuLysAspVal\*\*SerAspLysValLysLeuVal\*\*AspIleLeuAlaPheLeuAla 484  
 Db 2279 CAAATCGAGATGCTGACTAGAAAAGTGCTTCAGTATATGAAC----- 2320  
 QY 485 ProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAsp 504  
 Db 2321 -----AAATATCTCTGAT-----AATGCAAAATTTAA-- 2347  
 QY 505 GluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAsp 524  
 Db 2348 -----AAGATATTTGAGTCAGATATGAGAGAACGAAAGAGATAATTAC----- 2392  
 QY 525 ProArgAspIleThrSerAspGluGlyAspAla-----Tyr 536  
 Db 2393 -----GGAAGTTTAGAAAATGATGCTTTGAAAGGCTATTTTGAGAAATATTTC 2440  
 QY 537 ValThrProHisMetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGlu 556  
 Db 2441 CTTACACCATTTAATAAA-----ATTAAGCAGATTTGTAGATGATTTGGAT 2485  
 QY 557 ArgAlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProSerThrAspHis 576  
 Db 2486 AAAAAGTAGACAA-----GATCAGCAGCACCAATTCGGAATTCAGAA 2533  
 QY 577 GlnAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAla 596  
 Db 2534 ATGGATCAGGCTAAGGAAAGGCTAAGATTGCTGTATCGAAGTAT----- 2578  
 QY 597 AlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLys 616  
 Db 2579 ATGAGTAAGGTT--TTAGATGGAGTTTCATCAACATCTGCAG-----AAGAAA 2623  
 QY 617 AsnGlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGlnTrpPhe 636  
 Db 2624 AATAACAGTAAAATTTGT-----GATCTTTT 2650  
 QY 637 AspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrVal 656  
 Db 2651 AAGGAA-----CTTGAAGCGATTAAACAACAATATTTTGTAT----- 2689  
 QY 657 LysTyrTyrValGluHisProAsnGluArgProHisSerAspAsn-----Gly 672  
 Db 2690 -----ATTGCAATGCAAGACTGAAGTAGAGATTGATAACTTAGTAGTACACGATGCA 2740  
 QY 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692  
 Db 2741 TTTCAAAAATGATGCTACTGTTGTTGTAATTTTCAAAAAGGT---CTAGAGACAAATACG 2797  
 QY 693 ThrGluLysProSerGluGluLys-----ProGlnThrGluLysProGluGlu 708  
 Db 2798 CCAAAAATCTCAGATACACCGAAGATTCCAGAGGCTTCAAGCCCCAGATACACCGGAG 2857

QY 709 GluThrProArgGluGluLysProGlnSerGluLys-----Pro 721  
 Db 2858 GCTCCAGACACCGCATGTTCCGGAATCACCAGAGCCCGCAGACCGCGTGTCCG 2917  
 QY 722 GluSerProLysProThrGluGluProGlu-----GluSerProGluGluSerGluGlu 739  
 Db 2918 GAATCACCAGAACTCCAGAGACGCCATGTTCCGGAATCACCAGAGCCCGCAGACGA 2977  
 QY 740 ProGlnVal---GluThrGluLysValGluGlu 749  
 Db 2978 CCGGTGTTCCGGAATCACCAGACTCCAGAA 3010

## RESULT 23

US-08-714-481-1  
 ; Sequence 1, Application US/08714481  
 ; Patent No. 5766606  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brady, L. Jeannine  
 ; TITLE OF INVENTION: Cloning of No. 5766606-IGA Fc Binding Forms of  
 ; TITLE OF INVENTION: the Group B Streptococcal Beta Antigens  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Saliwanchik & Saliwanchik  
 ; STREET: 2421 N.W. 41st Street, Suite A-1  
 ; CITY: Gainesville  
 ; STATE: FL  
 ; COUNTRY: USA  
 ; ZIP: 32606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA: US/08/714,481  
 ; APPLICATION NUMBER: US/08/714,481  
 ; FILING DATE: 16-SEP-1996  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/242,932  
 ; FILING DATE: 16-MAY-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Saliwanchik, David R.  
 ; REGISTRATION NUMBER: 31,794  
 ; REFERENCE/DOCKET NUMBER: UF142  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 904-375-8100  
 ; TELEFAX: 904-372-5800  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4200 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; US-08-714-481-1

Alignment Scores:  
 Pred. No.: 3,73e-05 Length: 4200  
 Score: 163.00 Matches: 166  
 Percent Similarity: 35.26% Conservative: 127  
 Best Local Similarity: 19.98% Mismatches: 316  
 Query Match: 4.06% Indels: 222  
 DB: 1 Gaps: 44

US-09-765-272A-66 (1-763) x US-08-714-481-1 (1-4200)

QY 8 HisGlnAlaGlyGlnValLysLys-----GluSerAsnArgValSerTyrIleAspGly 25  
 Db 917 CATCAAAAAGTCAAGTTGAAAAAATGCGAGAGCAAAAGGAATTCACAAATGAAGATAAA 976  
 QY 26 Asp-----GlnAlaGlyGlnLysAlaGluAsnLeu 35

Db 977 GATTCCTAGCTGAAAAAATCGAAGATATTCGTAAACAAGCTCAACAAGCAGATAAAAA 1036  
 QY 36 ThrProAspGluValSerLysArgGlu----- 44  
 Db 1037 GAAGATGCCGAAGTAAAGGTTCTGTGAAGAACTAGGTAAACTCTTTAGTCAACTAAAGCT 1096  
 QY 45 GlyIleAsnAlaGlu-----Gln\*\*\*ValIleLysIleThrAspGlnGlyTyrValThr 62  
 Db 1097 GGTCTGGATCAAGAAATTCAGAGCATGTGAAGAAACGAGTAGTAGGAGAAATACT 1156  
 QY 63 SerHisGlyAspHisTyrHisTyrAsn-----GlyLys 74  
 Db 1157 CAGAAAGTTGAGAA---CACTATGCTATATGCTTCAGAACCTGCTCAAAAATCTCTT 1213  
 QY 75 ValProTyrAspAlaIleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeu 94  
 Db 1214 GAAGAACTAGATAAGGCACTACCAATGAACAAGCTACACAAGTTAAAATCAATCTTA 1273  
 QY 95 LysAspSerAspIleValAsnGluLysGlyTyrValIleLysValAsnGlyLys 114  
 Db 1274 GAAAAAGCTCAAAAGCTCAAGAAATACAA---CCTCTTATCAAGAAACGAATGTGAAA 1330  
 QY 115 TyrTyrValTyrLeuLysAspAlaIleHisAlaAspAsnIleArgThrLysGluIle 134  
 Db 1331 TTGTATAAGGCTATG-----AGTGAGAGCTG 1357  
 QY 135 LysArgGlnLysGlnGluArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAla 154  
 Db 1358 GAGCAGGTTGAGAGGAATTAACAATATTCGGAAGCTTAATTTAGAGATTTGGTTCG 1417  
 QY 155 AlaAlaArgAlaGlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAsp 174  
 Db 1418 AAATCTAAA-----GAA 1429  
 QY 175 IleIleGluAspThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIle 194  
 Db 1430 ATCGTAAGAGAA-----TACGAAGGAAAA 1453  
 QY 195 ProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLys 214  
 Db 1454 CTTAATCAATCTAAAAATCTTCCAGAAATTAAGCAACTAGAGAGGAAGCTCAATCGAAG 1513  
 QY 215 -----GlnGlySerArgProSerSerSerSerSerSerSerSerSerSerSerSer 229  
 Db 1514 TTGAACAAGTTGTGGAGGATTTTAAAAAATTTAAAAACGTGACAGCAAGTGACACCA 1573  
 QY 230 -----AlaGlnProArgLeuSerGluAsnHisAsn-----Leu 240  
 Db 1574 AAAAAACGTGTCAACAGAGATTTAGCTGCTATGAAAAATTAATCAACAAAAAGATTGAGTTA 1633  
 QY 241 ThrValThrProThrTyrHisGlnAsn-----GlnGlyGluAsnIleSerSer 256  
 Db 1634 ACAGTTTCAACCA-----GAGAAATATCATGTATATGAGGTGAAGCGGAAATTT 1684  
 QY 257 LeuLeuArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeu 276  
 Db 1685 ACAGTCACA-----GCTAAAAAGTGTTCGAGACCGCTGGACTTCAGTGATCTT 1735  
 QY 277 Ile-----PheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValPro 294  
 Db 1736 TTAACAAAATATAATCCGTCT---GTATCAGATAGAAATTTAGTACAAATATAAGACTAAC 1792  
 QY 295 HisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIle 314  
 Db 1793 ACGGATAATCATAAG-----ATTCCGAAATTCATCTATCAAGAAATTTGAAGCTAAATGAA 1846  
 QY 315 AlaAlaIleIleProLeuArgTyrArgSerAsn-----His 326  
 Db 1847 AGTCAACAGTGACTCTAAAGCTAAAGATGATCTTGGCAATGTAGTGTGAAAAAACATTC 1906  
 QY 327 TrpValProAspSerArgProGlnGlnProSerProGlnSerThrProGluProSerPro 346  
 Db 1907 ACTATTACAGTGCAAAAGAGAGAGAGAAACAAGTTCTTAAACACACGAGCAGAAAGAT 1966

```
347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLeu 366
1967 TCTAAACCGAAGAAAGGTTCTCCTCAAGAACCAAAATCAAT-----GACAAAGATCAA 2020
367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386
2021 TTCAAGAGTTCATTAAATCAGCTCAACAGAACTCGAAAGTGTAGAAAAGCAATAAAA 2080
387 ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406
2081 GAATTAATCGGACCAACAGAGATTCATCCATCCAGAGTATGTTATTCAAAATCTATT 2140
407 -----AlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuPro 424
2141 TGGGAGTCAAAAAGAGCTCTCAGGAAGGCATACCAAGTTTAAAGAGATTATTGGT 2200
425 SerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAsp 444
2201 GATTTCATCTCAAAATPACTACACAGACACTAT----- 2233
445 LeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArg 464
2234 -----TTTAAACAATATAAATCT--GATTT-----ATGAATTATCACTTCATGCA 2278
465 LeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeuAlaPheLeuAla 484
2279 CAAATGGAGATCTGACTAGAAAAGTGGTTCAGTATATGAAC----- 2320
485 ProIleArgHisProGluArgLeuGlyLysProAsnAlaGluIleThrTyrThrAspAsp 504
2321 -----AAATATCCCTGAT-----AATGCAGAAATATAA----- 2347
505 GluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAsp 524
2348 -----AGATATTGAGTCAGATATGAGAGACGAAAGAGATTAATAC----- 2392
525 ProArgAspIleThrSerAspGluGlyAspAla----- 2392
2393 -----GGAAGTTTAGAATAATGATGCTTTGAAAGGCTATTTTCAGAAATATTTC 2440
537 ValThrProHisMetThrHisSerHisThrLysLysAspSerLeuSerGluAlaGlu 556
2441 CTTACACCATTTAATAA-----ATTAAGCAGATGTAGATGATTGGAT 2485
557 ArgAlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHis 576
2486 AAAAAAGTAGAACAA-----GATCAGCCAGCACCAATCCGAAATTCAGAA 2533
577 GlnAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAla 596
2534 ATGGATCAGGCTAAGGAAAGGCTAAGATTGCTGTATCGAAGTAT----- 2578
597 AlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLys 616
2579 ATGAGTAAGGTT--TTAGATGGAGTTTCATCAACATCTCAG-----AAGAAA 2623
617 AsnGlySerLeuIleProHisTyrAspHisTyrHisAsnLysPheGluTyrPhe 636
2624 AATAACAGTAATAATGTT-----GATCTTTT 2650
637 AspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrVal 656
2651 AAGGAA-----CTTGAAGCGATTAAACAACTATTTTGTAT----- 2689
657 LysTyrTyrValGluHisProAsnGluArgProHisSerAspAsn-----Gly 672
2690 -----ATTGACAATGCAAGACTGAAGTAGAGATTGATACTTAGTACACGATGCA 2740
673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692
2741 TTCTCAAAAATGAATGCTACTGTTCTTAATTTTCAAAAGGT-----CTAGAGACAAATACG 2797
```

```
QY 693 ThrGluLysProSerGluGluLys-----ProGlnThrGluLysProGluGlu 708
Db 2798 CCGAAACTCCAGATACACCGAAGATTCCAGAGTACCTCAAGCCCGCATACACGCGAG 2857
QY 709 GluThrProArgGluGluLysProGlnSerGluLys-----Pro 721
Db 2858 GCTCCAGACACACCGCATGTTCCGGAATCACCAGGCCCCCAGAGCCCGGTGTTCCG 2917
QY 722 GluSerProLysProThrGluGluProGlu-----GluSerProGluGluSerGluGlu 739
Db 2918 GAATCACCAGAAAGACTCCAGAGCACCAGCATGTTCCGGAATCACCAGGCCCCCAGAGCA 2977
QY 740 ProGlnVal---GluThrGluLysValGluGlu 749
Db 2978 CCGCGTGTTCGGAATCACCAGAGACTCCAGAA 3010

RESULT 24
US-08-923-992A-1
; Sequence 1, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF INVENTIONS: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4200 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 320..3811
US-08-923-992A-1
```

```
Alignment Scores:
Pred. No.: 3,73e-05 Length: 4200
Score: 163.00 Matches: 166
Percent Similarity: 35.26% Conservative: 127
Best Local Similarity: 19.98% Mismatches: 316
Query Match: 4.06% Indels: 222
DB: 3 Gaps: 44
```

US-09-765-272A-66 (1-763) x US-08-923-992A-1 (1-4200)



```

b 2690 -----ATTGCAATGCAAGACTGAAGTAGAGATTGATACTTAGTACACCGATGCA 2740
      |||||
y 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692
      |||||
b 2741 TTCTCAAAATGAAGTACTGTTGCTAATATTCAAAAGGT---CTAGAGACAATACG 2797
      |||||
y 693 ThrGluLysProSerGluGluLys-----ProGlnThrGluLysProGluGlu 708
      |||||
b 2798 CCAGAACTCCAGATACACCGAAGATTCCAGAGCTACCTCAAGCCCCAGATACACCGCAG 2857
      |||||
y 709 GluThrProArgGluGluLysProGlnSerGluLys-----Pro 721
      |||||
b 2858 GCTCCAGACACACCGCATGTTCCGGAATCATCCAAAGGCCCCAGAGCACCGCGTGTCCG 2917
      |||||
y 722 GluSerProLysProThrGluGluProGlu-----GluSerProGluGluSerGluGlu 739
      |||||
b 2918 GAATCACCAAGACTCCAGAGACCGCATGTTCCGGAATCACCAAGGCCCCAGAGCA 2977
      |||||
y 740 ProGlnVal---GluThrGluLysValGluGlu 749
      |||||
b 2978 CCGCGTGTTCGGGAATCACCAAGACTCCAGAA 3010
      |||||

```

RESULT 25  
 CT-US95-06111-1  
 Sequence 1, Application PC/TUS9506111  
 GENERAL INFORMATION:  
 APPLICANT: Street address: 186 Grinter Hall  
 APPLICANT: City: Gainesville  
 APPLICANT: State/Province: Florida  
 APPLICANT: Country: US  
 APPLICANT: Postal code/Zip: 32611  
 APPLICANT: Phone number: 904-392-8929  
 APPLICANT: Fax number: 904-392-6600  
 APPLICANT: Telex number:  
 TITLE OF INVENTION: Cloning of Non-IgA Fc Binding Forms of  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Saliwanchik & Saliwanchik  
 STREET: 2421 N.W. 41st Street, Suite A-1  
 CITY: Gainesville  
 STATE: FL  
 COUNTRY: USA  
 ZIP: 32606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/06111  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/242,932  
 FILING DATE: 16-MAY-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Saliwanchik, David R.  
 REGISTRATION NUMBER: 31,794  
 REFERENCE/DOCKET NUMBER: UFI42  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 904-375-8100  
 TELEFAX: 904-372-5800  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4200 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 CT-US95-06111-1

Alignment Scores: 3.73e-05 Length: 4200  
 Pred. No.: 163.00 Matches: 166  
 Score: 163.00  
 Percent Similarity: 35.26% Conservative: 127  
 Best Local Similarity: 19.98% Mismatches: 316  
 Query Match: 4.06% Indels: 222  
 DB: 5 Gaps: 44

US-09-765-272A-66 (1-763) x PCT-US95-06111-1 (1-4200)  
 QY 8 HisGlnAlaGlyGlnValLysLys-----GluSerAsnArgValSerTyrIleAspGly 25  
 |||||  
 Db 917 CATCAAAAAGTCAAGTTCGAAAAATGCGCAGAGCAAAAGGGAATCACAAAATGAAGATAAA 976  
 |||||  
 QY 26 Asp-----GlnAlaGlyGlnLysAlaGluAsnLeu 35  
 |||||  
 Db 977 GATTCTATCTGCTGAAAAAATCGAAGATATTCGTAACAAGCTCAACAAGCAGATAAAAA 1036  
 |||||  
 QY 36 ThrProAspGluValSerLysArgGlu----- 44  
 |||||  
 Db 1037 GAAGATGCCGAAGTAAGGTTCGTGAAGAACTAGGTAAACTCTTTAGTTCACCTAAAGCT 1096  
 |||||  
 QY 45 GlyIleAsnAlaGlu-----Gln\*\*\*ValIleLysIleThrAspGlnGlyTyrValThr 62  
 |||||  
 Db 1097 GGTCGTGATCAAGAAATTCGAAGCATGTGGAAGAAAGAAACGAGTAGTAGAGAAAATACT 1156  
 |||||  
 QY 63 SerHisGlyAspHisTyrHisTyrTyrAsn-----GlyLys 74  
 |||||  
 Db 1157 CAGAAAGTTGATGAA---CACTATGCTAATAGCTTCAGAACCTTGCTCAAAAATCTCTT 1213  
 |||||  
 QY 75 ValProTyrAspAlaIleIleSerGluGluLeuMetLysAspProAsnTyrGlnLeu 94  
 |||||  
 Db 1214 GAAGAACTAGATAAGCAACTACCAATGAACAAGCTACACAGTTAAAAATCAATCTTTA 1273  
 |||||  
 QY 95 LysAspSerAspIleValAsnGluIleLysGlyTyrValIleLysValAsnGlyLys 114  
 |||||  
 Db 1274 GAAACGCTCAAAAGCTCAAGAAATACAA---CCTCTTATCAAGAAACCAAGTGTGAA 1330  
 |||||  
 QY 115 TyrTyrValTyrLeuLysAspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIle 134  
 |||||  
 Db 1331 TTGTATAGGCTATG-----AGTGAGAGCTTG 1357  
 |||||  
 QY 135 LysArgGlnLysGlnGluArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAla 154  
 |||||  
 Db 1358 GAGCAGTTGAGAGGAATTTAAACATAATTCGGAAGCTAATTTAGAAGATTGCTGCG 1417  
 |||||  
 QY 155 AlaAlaArgAlaGlnGlyArgTyrThrAspAspGlyTyrIlePheAsnAlaSerAsp 174  
 |||||  
 Db 1418 AAATCTAAA-----GAA 1429  
 |||||  
 QY 175 IleIleGluAspThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIle 194  
 |||||  
 Db 1430 ATCGTAAGAGAA-----TACGAAGGAGAAA 1453  
 |||||  
 QY 195 ProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLys 214  
 |||||  
 Db 1454 CTTAATCAATCTAAAAATCTTCCAGAAATTAAGCAACTAGAGAGGAGGCTCATTCGAAG 1513  
 |||||  
 QY 215 -----GlnGlySerArgProSerSerSerSerSerSerSerSerSerSerSer 229  
 |||||  
 Db 1514 TTGAACAAGTTGTGGAGGATTTTAAAAAAATTTAAAAAGTCAGAGCAGTACACCA 1573  
 |||||  
 QY 230 -----AlaGlnProArgLeuSerGluAsnHisAsn-----Leu 240  
 |||||  
 Db 1574 AAAAAACGTTCAACACGAGATTTAGCTGCTAATGAAAAATTAATCAACAAAAGATTGAGTTA 1633  
 |||||  
 QY 241 ThrValThrProThrTyrHisGlnAsn-----GlnGlyGluAsnIleSerSer 256  
 |||||  
 Db 1634 ACAGTTTCACCA-----GAGAATATCACTGCTATATGAAGTGAAGACGTGAAATTT 1684  
 |||||  
 QY 257 LeuLeuArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeu 276  
 |||||

1685 ACAGTCACAC-----GCTAAAGATGATTGCAAGACGAGCTTGGACTTCAGTGATCTT 1735  
277 Ile-----PheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValPro 294  
1736 TTAACAATATATATCGCTCT-----GATCAGATAGAAATAGTACAAATTTAAGACTAAC 1792  
295 HisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGlyArgile 314  
1793 ACGGATAATCATAG-----ATTGCCGAATCATCTATCAAGAATTTGAAGCTAAATGAA 1846  
315 AlaArgIlelleProLeuArgTyrArgSerAsn-----His 326  
1847 AGTCAAAACAGTACTCTAAAGCTTAAAGATGATCTCGCAATGTAGTTGAAAAACATTC 1906  
327 TrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSerPro 346  
1907 ACTATTACAGTGCAGAAAGAGGAGGAAACAAGTCTCTAAACACACAGACGAGAAAGAT 1966  
347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366  
1967 TCTAABACGGAAGAAAGGTTCTCTCAAGAACCAAAATCAAT-----GACAAATCAA 2020  
367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386  
2021 TTACAAGAGTTGATTAAATCAGCTCAACAAGAACTGGAAGAGTTAGAAAAGCAATAAAA 2080  
387 ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406  
2081 GAATTAATGGAGCAACGACAGATTCCTCAATCCAGAGTATGATTTCAAAATCTATT 2140  
407 -----AlaLysGlnGluSerHisLysLeuGlyAlaLysLysThrAspLeuPro 424  
2141 TGGAGGTCAAAAAGAGCCTATCCAGAGACCAATACAGAGTTTAAAGAGATTATGGT 2200  
425 SerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAsp 444  
2201 GATTCATCTTCAAAATACTACAGACGACTAT----- 2233  
445 LeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArg 464  
2234 -----TTTAACAATAATAATCT-----GATTTT-----ATGAATATCACTTCATGCA 2278  
465 LeuLysAspVal\*\*SerAspLysValLysLeuVal\*\*AspIleLeuAlaPheLeuAla 484  
2279 CAATGAGATGCTGACTAGAAAGTGGTTCAGTATATGAC----- 2320  
485 ProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAsp 504  
2321 -----AAATATCTGTAT-----AATGCAGAAATTAAT----- 2347  
505 GluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAsp 524  
2348 -----AAGATATTGTAGTCAGATATGAGAGACGAAAGAAATAATTAC----- 2392  
525 ProArgAspIleThrSerAspGluGlyAspAla-----Tyr 536  
2393 -----GGAAGTTAGAAATATGCTTTGAAAGCTATTTTGAGAAATATTTTC 2440  
537 ValThrProHisMetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGlu 556  
2441 CTTACACCAATTAATAA-----ATTAAGCAGATGTAGATGATTTGGAT 2485  
557 ArgAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHis 576  
2486 AAAAAAGTAGAACAA-----GATCAGCAGCAGCACCAATTCGGAATAATTCAGAA 2533  
577 GluAspSerGlyAsnThrGluAlaGlyAlaGluAlaIleTyrAsnArgValLysAla 596  
2534 ATGGATCAGGCTAAGGAAAGGCTTAAAGTGTCTGTATCGAAGTAT----- 2578  
597 AlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLys 616  
2579 ATGAGTAAAGTT-----TTAGATGGAGTTCAATCAACATCTGCAG-----AAGAA 2623

QY 617 AsnGlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPhe 636  
Db 2624 AATAACAGTAAATGTT-----GATCTTTT 2650  
QY 637 AspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrVal 656  
Db 2651 AAGGAA-----CTTGAAGCGATTAAACAACAACAACTAATTTTCAT----- 2689  
QY 657 LysTyrTyrValGluHisProAsnGluArgProHisSerAspAsn-----Gly 672  
Db 2690 -----ATTGACAATGCAAAAGACTGAAGTAGAGATTGATACTTAGTACACGATGCA 2740  
QY 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692  
Db 2741 TTTCTAAAATAAGTACTGTTCTGCTAAATTTCAAAAGGT-----CTAGACACAATACG 2797  
QY 693 ThrGluLysProSerGluGluLys-----ProGlnThrGluLysProGluGlu 708  
Db 2798 CCAGAAACTCCAGATACACCGAAGATTCCAGAGCTACCTCAAGCCCCCAGATACACCGCAG 2857  
QY 709 GluThrProArgGluGluLysProGlnSerGluLys-----Pro 721  
Db 2858 GCTCCAGACACACCGCATGTTCCGGAATCACCAGAGCCGCCAGAGCCCGCGTGTTCG 2917  
QY 722 GluSerProLysProThrGluGluProGlu-----GluSerProGluGluSerGluGlu 739  
Db 2918 CAATCACCAGAGACTCCAGAGACCCGCGATGTTCCGGAATCACCAGAGCCGCCAGAGCA 2977  
QY 740 ProGlnVal---GluThrGluLysValGluGlu 749  
Db 2978 CCGCGTGTTCGGAATCACCAGAGCTCCAGAA 3010

## RESULT 26

US-08-956-171E-39  
; Sequence 39, Application US/08956171E  
; Patent No. 6593114  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; Gil H. Choi  
; Patrick S. Dillon  
; Craig A. Rosen  
; Steven C. Barash  
; Michael R. Fannon  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5256  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/956,171E  
; FILING DATE: 20-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/009,861  
; FILING DATE: January 5, 1996  
; APPLICATION NUMBER: 08/781,986  
; FILING DATE: January 3, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mark J. Hyman  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB248P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (240) 314-1224



TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 4522 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 39:

US-08-956-171E-39

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Conservative:

Mismatch:

Indels:

Gaps:

DB:

US-09-765-272A-66 (1-763) x US-08-956-171E-39 (1-4522)

2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21  
 1220 AGCTATGAATGGCAACATTTTGTGATA-ATTAAGAGAGTAGTGAATATGAGC 1278  
 22 TyrIleAsp-----GlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAsp 38  
 1279 TGGTTTGATAAATATTTCGCGAAGAT-----AATGATTCAAATGATGAC 1323  
 39 GluValSerLysArgGluGlyIleAsnAlaGluGln\*\*ValIleLysIleThrAspGln 58  
 1324 TTGATTTCATAGAAAGAAAAAGACGTCAGAAATCACAAAATATAGATACGATCATGAC 1383  
 59 GlyTyrValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAsp 78  
 1384 TCATTACTGCCTCAAAATATGATATTATAGTCGCGAGGGGAAATTCGCTTTTCCCT 1443  
 79 AlaIleIleSerGluGluLeuMetLysAspProAsnTyrGlnLeuLysAspSerAsp 98  
 1444 ATGACGTAGCTTATGAA-----AATGAAATGTTGACAACTCT--- 1482  
 99 IleValAsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyr 118  
 1482 ----- 1482  
 119 LeuLysAspAlaAlaHisAlaAspAsnIleArg---ThrLysGluGluIleLysArgGln 137  
 1483 -----GCAGATACTATTTCAGATGAAAAAGAACAAATACCATCGAGAC 1524  
 138 LysGlnGluArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaArg 157  
 1525 TATCGCAACAAAGCCAGATTCTCGTTCACAAAACGACAT-----CGC 1569  
 158 AlaGlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGlu 177  
 1570 CGTAGAAGAAATCAAACTCAAGAA-----CAAAATATATAGTGA 1611  
 178 AspThrGlyAspAlaTyrIle-----ValProHisGlyAspHisTyrHisTyr 193  
 1612 CAACGTGGGAATTTCTAAAATATATACACGAAAGTATATAAATATAGATCATTCACATTAC 1671  
 194 -----IleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyr 210  
 1672 CATACGAATAACCGGTACATATGTTTCTGCATTT----- 1707  
 211 TrpAsnGly-----LysGlnGlySerArgProSerSerSerSerTyrAsnAlaAsn 228  
 1708 ---AATGGTATTGAGAGAAAGCCACAGCCAAAACACATATATGTTTCTTAATAT 1764  
 229 ProAlaGlnProArgLeuSerGluAsnHisAsnLeu---ThrValThrProThrTyrHis 247  
 1765 -----ACAAATCATCGTGTAAAGATTCAAAGATTCACGATTCATCAC 1803  
 248 GlnAsnGln-----GlyGluAsnIleSerSerSerLeuLeuArgGluLeuTyrAlaLys 264

1804 AAGAAAAGTTTCAAGACTTCAGAGGTACCGTCAGCTATTTTGGCACAATGAACCTAAA 1863  
 265 ProLeu-----SerGluArgHisValGluSerAsp 274  
 1864 AAGTTAGAAAATGGTGTATCCCTGTAAAGTAACCTTCAGAAAAA---GTTGAGTCAGAT 1920  
 275 GlyLeuIlePheAspProAlaGlnIleThrSerArgThrAla----- 288  
 1921 AAACAAAAATATGATAAATATGTAGCTAAGCAGCAACGCTCTCAAAATAAACAATTAGAA 1980  
 289 -----ArgGlyValAlaValProHisGlyAsnHis 298  
 1981 CAAGAAAAACAATAATGATAGTTGTCTCAAAACAAGAACTGCATCTAAATCATCTGAT--- 2037  
 299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArg----- 316  
 2038 -----GAAATGTATCATCAACAACAATAATCAATGCTTAATTATTCA 2079  
 317 -----IleIleProLeu 320  
 2080 AAGTTTCATTAATACTATCAAAATTTGAAAATATTTATGCTTCACAAATGTTTGAAGAAAT 2139  
 321 ArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSer 340  
 2140 AGACGTCAACGAGAACCGTAAAGTGTCTCAAAAGCGTTCGATTATAAAAAAGCGTTGCAACAA 2199  
 341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360  
 2200 AAGCGTCAAGAACATNAAAAACGAAGAG-----CAAGATGCAATACAACT 2244  
 361 ProfileAspGluLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377  
 2245 GCATTGTGAAATGTATGTCTAAACAGCGAAGCTATGTTGGTGATAGTTCAATTAAT 2304  
 377 ----- 377  
 2305 GATGATAGTGACTTAAACAGATAATAGTACAGATGCTAGTCAGCTTCATACAAATGCGATA 2364  
 377 ----- 377  
 2365 GAGAAATGAAACTGTATCAAAATGATGAAAAATAAACAAGCGTCAATACAAAATGAACACT 2424  
 378 -----TyrValPheGluGluAsnGlyValSerArgTyr 388  
 2425 AATGACACTCATGTAGATGAAGTCCATCAATATAGAGGAAGTTAGTTTGAATCAAGTA 2484  
 389 IleProAlaLysAspLeuSerAlaGlu-----ThrAlaAlaGlyIleAspSerLysLeu 406  
 2485 TCGACACAAAAACAATTTGTTCAGATGATGAAGTTACGGTTTCGAAATGTAACGCTCTCAA--- 2541  
 407 AlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys---LysThrAspLeuProSer 425  
 2542 ---CATCAATCAGCACTACACATCAAGCTTGAAGTAAATGATGAAGATGAACATAAAAA 2598  
 426 SerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeu 445  
 2599 CAATCCAGA-----TTAATTGCTGATTTCAGAGAAGATGGA 2634  
 446 LeuAspAsnLys-----GlyArgGlnValAsp-----PheGluAla 457  
 2635 GCAACGATTAAGAGATATTCAGKAAGTCAATCGATGATGATGATGATGATGATGATGATGAT 2694  
 458 LeuAspAsnLeuLeuGluArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\* 477  
 2695 AATGATACA-----GAATAGATGAGGATACCTACTCTCAAAATATCGAA 2736  
 478 AspIleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAla 497  
 2737 GATAATACCAATAGAAACGCGTCTGAAATGCAAT-----GTAGACGCTCTCTAAAAACG 2787  
 498 GlnIleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThr 517





EARLIER APPLICATION NUMBER: US 09/36987

NUMBER OF SEQ ID NOS: 39

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 80161

TYPE: DNA

ORGANISM: Saccharopolyspora spinosa

US-09-370-700-1

Alignment Scores:

SeqScore: 0.00896 Length: 80161

Matches: 159.50 Matches: 158

Conservative: 30.55% Conservative: 87

Mismatches: 19.70% Mismatches: 284

Indels: 3.97% Indels: 275

Gaps: 3 Gaps: 27

US-09-765-272A-66 (1-763) x US-09-370-700-1 (1-80161)

```

QY 187 HisGlyAspHisTyrHisTyrileProLysAsnGluLeu-----Ser 200
Db 44150 CACGTCCTCCCGACCAACCGCGTCCGACCAATCCCTCGCGCGAGCTGCTCGGCAAC 44091
QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyr----- 210
Db 44090 GCATCCAGTAGCGGTTCGACGCCGATACGACACCTGTCCGCGGCTCCCGACACCC 44031
QY 211 -----TrpAsnGlyLysGlnGlySerArgProSerSerSerSerSer 224
Db 44030 GACACCGAGGAGAACATCACAAACGCATCCAGTCCGAGTCCGCGCAGCAACTCGTCCAGA 43971
QY 225 TyrAsnAlaAsnProAlaGlnProArg--LeuSerGluAsnHisAsnLeuThrValThrP 244
Db 43970 TTC-----GCGCACCAACGACCTTGGCGGACACACCGCTGCAAAATCCGCG 43923
QY 244 roThr-----TyrHisGlnAsnG 250
Db 43922 AAACCATCTCCACTACGTGCGCGCGCTCGCGACCCCGCGGCGGTATACCGCGAGTC 43863
QY 250 In----- 250
Db 43862 AGCGGACACTCCCGACCGATCCCTGCCAACACTCCAGCCACTCGTCAGCATCAGCCACG 43803
QY 251 ----GlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysProLeu-----S 267
Db 43802 TCGAGGGGACATCGACACCGGAGCACCAGCGGCTCCAGCTCCGCGCGCAATTCGCCA 43743
QY 267 exGluArgHisValGluSer--AspGlyLeuIlePheAspProAlaGlnIleThrSerAr 286
Db 43742 GCGCGGAGCGTCCGCTCTCGACGGTGGTCAACACACAGCGTGTCTCAGCCCGGACCG 43683
QY 286 gThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyr----- 304
Db 43682 GCCAACACCGCGGCAACATGCGGCGCCCAACCGCCCGTACCAACCGTCCACCAACCGTC 43623
QY 305 ----GluGlnMetSerGluLeuLysArgIleAlaArgIleIleProLeuArgTyrAr 323
Db 43622 CCGCGGAGGCGCACACCGACCGGACCGCGACCGCGACCGCGGTTTCCAACCGCGGCA 43563
QY 323 gSerAsnHis-----TyrValProAspSerArgProGluGlnProSerProGlnSe 340
Db 43562 AACACACACCGCGGATGGGACTTGTCTCGCCACCATCCGCGCAACACACCGCGC 43503
QY 340 rThrProGluProSerProSerProGlnPro----- 350
Db 43502 AACCGGCGCCCAATCGGTCTGCCAGTCGATGTCGAGGTGATCAAGCCACCGCAGAAC 43443
QY 351 -----AlaProAsnProGlnProAlaProSer-A 360
Db 43442 TCGGGGTGCTCCAGTCCCGGACACGACGACGCGCGCGCGCGCGGCTCTCGGAGAC 43383
QY 360 snProfileAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValP 380

```

```

Db 43382 AC-----GCTACATCT 43371
QY 380 heGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaA 400
Db 43370 TCCACGCGACCGCGCGCGCTCGCGCCCAAAAGCGCTCGAGCTCGCGCTTCGCGC 43311
QY 400 laGlyIleAspSerLysLeuAlaLysGlnGluSerLysHisLysLeuGlyAlaLysL 420
Db 43310 AAGGC---CCGCAAGAGCTGCACAGTCAGCGCGAAACAGTGTGCGAGCAGGAGATCCC 43254
QY 420 yThrAsp-----LeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuL 438
Db 43253 GGAGCAGCGCGATTTCATCGAGCCCAACAGCGAAAGACCGCCAGGTGTGACCGCATCC 43194
QY 438 euAlaArgIleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaL 458
Db 43193 GTCAGCGCATTCGCGCGCGCTGTCGCGAGCTCTCGCGGTCTCTCTTCGACAGG 43134
QY 458 eu-----AspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuV 476
Db 43133 CCGATCCGAGACCTCGG-CGCGCTCCGACCGAGTGAATCTCTCAAGCGC----- 43082
QY 476 al***AspIleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProA 496
Db 43081 -----TGACCAATCCGTCTATCCGACGACCGCGCGCAGCACAA 43045
QY 496 snAlaGlnIleThrTyrThrAspAspGluIleGlnValAlaLysLeuAla----- 512
Db 43044 CAGCCAGTACCGGAATCTTCGCGGCGCGCACCTCCGAGAGGAATCCCACTGATTCG 42985
QY 513 ----GlyLysTyrThrThrGluAspGly----- 520
Db 42984 GTACCGCGAGATTCGACGAGCGAACGTTTCGCGAGTTCGCGCAGGAGCAACGCG 42925
QY 520 ----- 520
Db 42924 CGGCAATACATCGTCAACGATGCGATGTGTCGTCGCCCAAGATCCAGAGCGCG 42865
QY 521 -----TyrIlePheAspProArgAspIleThrSerAspGluGlyAspA 535
Db 42864 GACCAACGTTCCGACATCGCTTGGTCAACCGCGTCCAGAACTGTCGCGCACCTCGCG 42805
QY 535 laTyrValThrProHisMetThrHisSerHisTyrIleLysLysAspSerLeu----- 552
Db 42804 GATGTCGCGGTCGCTGCTGCGCGGCTAGTCCGATCCAGTAGTGTGTCGCTG 42745
QY 553 -----SerGluAlaGluArgAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuT 570
Db 42744 GAAGGCATACGTTGGCAACTCGAGCTGCTAGCGCGGTACCGCGGCAAAACGCGTCGCA 42685
QY 570 hrProProSerThrAspHisGlnAspSerGlyAsnThr----- 582
Db 42684 GTCACCT-----CACACACCGGTTGTGAGATGCCCGAGGCTGTCTATCCGCAAG 42631
QY 583 -----GluAlaLysGlyAlaGluAlaIleTyrA 592
Db 42630 GGTCTCGTCGCTCCCTCGCATCATGATCGAGGAATCCGCCCAACCGCATCGA 42571
QY 592 snArgValLysAlaAlaLysLysValPro-----LeuAspArg 604
Db 42570 CTGGTCGACACTCCCGAACACAGTCGCGCGGCTTCGCTCCGCGCGCGAGTCCGACG 42511
QY 605 MetProTyrAsnLeuGlnTyr-----ThrValGluValLysAsnGly----- 618
Db 42510 CGTGCCTATGCTTCGCTCCAGGCGCTCGACACCGTTCGCGGAAACCGAGCGGTTCAG 42451
QY 619 -----SerLeuIleIleProHisTyrAspHisTyr----- 628
Db 42450 CGCTGGCGGACCAACATCCCTGTGTGCTCTATTCGCGGCTCTGTGAGCTCACCCTG 42391
QY 629 -----HisAsnIleLysPheGluThrPheAspGluGlyLeuTyrGluAlaProLysGly 646

```

Db 42390 CAACGTGCACACAAAGCGGTGCTGGCTTACCGAACTCCGACAGCTCGCGCATCTGGGC 42331  
2Y TyThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisProAsnGluArg 666  
Db 42330 GAATCGGCCCAACATCGGTTCATCGGTGCGCAATGAAAGATGGAGACCCGCAAGCG 42271  
2Y 667 ProHisSerAspAsnGlyPheGlyAsnAlaSerAsp- 678  
Db 42270 CCGCGTCCGAGCTCGAGTCTCGGTTCATCGCGCGCGCGAGATGCTGGCCAGCACCTCCCGATC 42211  
2Y 678 ----- 678  
Db 42210 GCCGACCAACACCAACCGAACCAGCGGCATTTAGCGCGCGCGAGCTCACGAGCTCTCTCCCG 42151  
2Y 679 -----HisValGlnArg 682  
Db 42150 ACCGCGCAGCAATGCGCGCAACAAAGTCTTCAACGACGACCGCAACATCG-CGCGGT 42092  
2Y 683 AsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGlu- 698  
Db 42091 CAGAGGCGAGGCTTGATCAACCGCGCGCGCGCGCCACCAACGAGCGCATCCGCA 42032  
2Y 699 -----GluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLys 715  
Db 42031 ACACAACACGCGCAGCGCAACCGCGCGCGCA-----ACTCCCGACCGAATGCC 41981  
2Y 716 ProGlnSerGlu-LysProGluSerProLysProThrGluGluGluProGluLysProGlu 735  
Db 41980 CCATCACCACATCCGCGCGCAACCGCGCGCA-----ACCCAGCGCGCGCGCGCTT 41924  
2Y 735 uGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArg 752  
Db 41923 GCAGCGCGAAGACAGACCGGACTGCGCGCCACAGGCTGTGATCCAGCAGCTGCGC 41872

## RESULT 29

US-09-603-207-1/c  
Sequence 1, Application US/09603207B  
Patent No. 6521406  
GENERAL INFORMATION:  
APPLICANT: Baltz, Richard H  
APPLICANT: Broughton, Mary C  
APPLICANT: Crawford, Kathryn P  
APPLICANT: Madduri, Krishnamurthy  
APPLICANT: Treadway, Patti J  
APPLICANT: Turner, Jan R  
APPLICANT: Waldron, Clive  
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide  
FILE REFERENCE: 50489 DIV1  
CURRENT APPLICATION NUMBER: US/09/603,207B  
CURRENT FILING DATE: 2000-06-23  
EARLIER APPLICATION NUMBER: 09/370,700  
EARLIER FILING DATE: 1998-03-09  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 1  
LENGTH: 80161  
TYPE: DNA  
ORGANISM: Saccharopolyspora spinosa  
US-09-603-207-1

Alignment Scores:  
Alignment No.: 0.00896 Length: 80161  
Score: 159.50 Matches: 158  
Percent Similarity: 30.55% Conservative: 87  
Best Local Similarity: 19.70% Mismatches: 284  
Query Match: 3.97% Indels: 275  
DB: 4 Gaps: 27

US-09-765-272A-66 (1-763) x US-09-603-207-1 (1-80161)

2Y 187 HisGlyAspHisTyrHisTyrTrpLeuProLysAsnGluLeu-----Ser 200  
Db 44150 CACGGTCCCAACGCAACCGCGGTTCGCAACCAATCCCTTCGCGCGAGCTGCTCGCGCAAC 44091

QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyr----- 210  
Db 44090 GCATCCAAGTAGGCGTTTCGACGCGCATACGACCCCTGTCGCGCGCTCCCCACACCCCC 44031  
QY 211 -----TrpAsnGlyLysGlnGlySerArgProSerSerSerSerSer 224  
Db 44030 GACACCGAGGAGAACATCAACAAACGATCCAGCTCCGAGTCGCGCAGCAACTCGTCCAGA 43971  
QY 225 TyrAsnAlaAsnProAlaGlnProArg--LeuSerGluAsnHisAsnLeuThrValThrP 244  
Db 43970 TTC-----GCCGACACCGACCTTCGCGGCAACACCGCTGCAAAATCCGCC 43923  
QY 244 rothr-----TyrHisGlnAsnG 250  
Db 43922 AACCCCATCTCCATAGTCGTCGCGCGCTCGCGACCCCGCGCGCTGTACCAACCGCAGTC 43863  
QY 250 ln----- 250  
Db 43862 AGCGGACATCCCCACCGATCCCTGCCAAACATCCAGCCACTGCGTCAGATCAGCCACG 43803  
QY 251 -----GlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysProLeu-----S 267  
Db 43802 TCGCAGGCGCAATCGACACCCGAGCACCCGAGCGCTCCAGCTCCGCGCGCAATTCGCCA 43743  
QY 267 erGluArgHisValGluSer--AspGlyLeuIlePheAspProAlaGlnIleThrSerAr 286  
Db 43742 CGCGCGCGAGCGTCTCTCTCTCGAGCGTGTCAACACCGCTGTCTAGCCCGCGCACCG 43683  
QY 286 gThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyr----- 304  
Db 43682 GCCAACACCGGCGCAACATGCGCGCCCAACCGCGCTACCAACCGCTACCAACACCGTC 43623  
QY 305 ----GluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrAr 323  
Db 43622 CCCGAGGACGCCACACCGACCCCGCGCACCGACCGCGCTTCCAAACCGCGCGCACG 43563  
QY 323 gSerAsnHis-----TrpValProAspSerArgProGluGlnProSerProGlnSe 340  
Db 43562 AACACACCCACCGCGGAATGGCGACTTGATCTCTCGCACCATCCGCGCAACACACCGCGC 43503  
QY 340 rThrProGluProSerProSerProGlnPro----- 350  
Db 43502 AACCGGCGACCCCATCGGTCTGCCAGTCCGATGCGAGCTCGATCAGCCACCCCGAGAC 43443  
QY 351 -----AlaProAsnProGlnProAlaProSer-A 360  
Db 43442 TCGCGGTGCTCCAGTCCCGCGCACGACGACGCGCGCTGCTCCGAGAC 43383  
QY 360 snProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValP 380  
Db 43382 AC-----GGGTACATCT 43371  
QY 380 heGluGluAsnGlyValSerArgTyrTrpLeuProAlaLysAspLeuSerAlaGluThrAlaA 400  
Db 43370 TCCAAACGCGAGCGCGCGGTCTACCGCCCAAAAGGCGCTCGACGTCGCGCTCCGC 43311  
QY 400 laGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysL 420  
Db 43310 AAGGC---CCGCAAGCTGCACAGTCAGCGCGAAACCCAGTTGGCAAGCAGGAGATCC 43254  
QY 420 ysThrAsp-----LeuProSerSerSerArgGluPheTyrAsnLysAlaTyrAspLeuL 438  
Db 43253 GGAGCGCGGATTCATCGAGCCCCCAACAGCGAAGAGCGCACCGCGAGTTGACCGCATCC 43194  
QY 438 euAlaArgIleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaL 458  
Db 43193 GTCAGCGATTTGGCGCGCTGTGCGACGTCCTCGCGGTGCGGTCTCTTCGCGACAGG 43134  
QY 458 eu-----AspAsnLeuLeuGluLeuLysAspVal\*\*\*SerAspLysValLysLeuV 476  
Db 43133 CCGATCCGAGCAGCTCGG-CGCGCGCTCCGACCACTCGTCAAGCG----- 43082

476 aI\*\*AspIleLeuAlaPheLeuAlaProfileArgHisProGluArgLeuGlyIysProA 496  
43081 -----TGACCAATCCGTCATCCGACGACCGGGCACCACAA 43045  
496 snAlaGlnIleThrTyrThrAspAspGluIleGlnValAlaIysLeuAla----- 512  
43044 CAGCCAGGTACCGGAATCTTCGGGGCGCGCACCTCCCGAGAGGAATGCCAACTGATTCG 42985  
513 -----GlyIysTyrThrThrGluAspGly----- 520  
42984 GTACCGGCAGGAATCGACGACGACGAGTTCGGGAGTCCGCTTCGCCAGGAGCAACAACG 42925  
520 ----- 520  
42924 CGGCAATACATCGCTCAACGATGCGCATGTGCTGCTCGCGCCCAAGATCAGACGGCG 42865  
521 -----TyrIlePheAspProArgAspIleThrSerAspGluGlyAspA 535  
42864 GACCAAGTTGCCAGATCGCTTGGTCAACCGCGGTCCAGAACTGTTCCGCCAACCTCGCC 42801  
535 laTyrValThrProHisMetThrHisSerHisTyrIleIysIysAspSerLeu----- 552  
42804 GATGTCGCGCGTGGCTGCTGGCGCGCACTGGACTCGACTGATGCTGTGCTGTTG 42745  
553 -----SerGluAlaGluArgAlaAlaGlnAlaTyrAlaIysGluIysGlyLeuT 570  
42744 GAAGGCATACGTTGGCACTCGAGCTGCTGCGCGGTACCGCGCAAAAACGCTGCCA 42685  
570 hrProProSerThrAspHisGlnAspSerGlyAsnThr----- 582  
42684 GTCCACCT-----CACCAACACGGGTGGAGATGCCGAGGGCTGTCTATCACCAGAACG 42631  
583 -----GluAlaIysGlyAlaGluAlaIleTyrA 592  
42630 GGTCTGTCGCGTCCCTCGGCATCAGTGGGATCGAGAAATCCGCCCAACCCGATCGGA 42571  
592 snArgValIysAlaAlaIysIysValPro-----LeuAspArg 604  
42570 CTCGTCGCACACTCCCGAAACCAAGTCCCGCAGCGTTCGTCGGGGCGAGCTCGACGAC 42511  
605 MetProTyrAsnLeuGlnTyr-----ThrValGluValIysAsnGly----- 618  
42510 CGTGCCTATGCGCTGCGTCCGAGCGCTGGACACCGTCCGGCAACGAGCGGGTTCAAG 42451  
619 -----SerLeuIleIleProHisTyrAspHisTyr----- 628  
42450 CGCCTGGCGCACCCCAATACCTGCTGCTCATTTCCGCGCTCTGTGAGCTCACCGGT 42391  
629 -----HisAsnIleIysPheGluTyrPheAspGluGlyLeuTyrGluAlaProIysGly 646  
42390 CAACGTCGACAAAGCGGTGCTGCTGGCTTACCGAACTCCGCAGACTCGCGGATCTGGGC 42331  
647 TyrThrLeuGluAspLeuLeuAlaThrValIysTyrTyrValGluHisProAsnGluArg 666  
42330 GAACTCGCCCAACATCGGTGCCATCGGTGCGAATGAAAGCATGGGAGACCCCGAACGG 42271  
667 ProHisSerAspAsnGlyPheGlyAsnAlaSerAsp----- 678  
42270 CGCGCTCGGATCTCGAGCTCGGTACGCGCGCGAGCATGTGCGCAGCACCTCCCGATC 42211  
678 ----- 678  
42210 GCCCGACAACACCAACCGAGCGGGCGATTTAGCGCGCGCAGCTCACGGACTCTCTCCCG 42151  
679 -----HisValGlnArg 682  
42150 ACCGGCCAGCAATGGCGCAACAGGTCTTCCACGACGACCCACCGCAACATCG-CGCCGT 42092  
683 AsnIysAsnGlyGlnAlaAspThrAsnGlnThrGluIysProSerGlu----- 698  
42091 CAGAGGGCAGGGCTTGATCAACCGGGCGCGCGGCCACCAACCGAGCCGATCCCGCA 42032  
699 -----GluIysProGlnThrGluIysProGluGluGluThrProArgGluGluIys 715

```

Db      42031  ACGACACAGCCAGCCGCAACAGCGCGGCCA-----ACTCCCGGACCGAATGCC 41981
QY      716   ProGlnSerGlu-LysProGluSerProLysProThrGluGluProGluGluSerProGln 735
Db      41980  CCATCACCACATCCGGCCGACAGCCGCCACGA-----ACCCAGCAGCCGCCAGAGCGCGGTT 41924
QY      735   uGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArg 752
Db      41923  GCAGCGCGAACAAGACCCGACTGCGCCCAAGCGTCTGATCCAGCAGCTGCCG 41872

RESULT 30
US-09-286-981B-29
; Sequence 29, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizemann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 29
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cDNA derived
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae
US-09-286-981B-29

Alignment Scores:
Pred. No.: 1,67e-05 Length: 1278
Score: 158.00 Matches: 128
Percent Similarity: 36.03% Conservat: 77
Best Local Similarity: 22.50% Mismatches: 183
Query Match: 3.93% Indels: 181
DB: 4 Gaps: 29

US-09-765-272A-66 (1-763) x US-09-286-981B-29 (1-1278)
QY      216   GlySerArgProSerSer-SerSerTyrAsnAlaAsnProAlaGlnProArgLeuSe 235
Db      12   GGTATACCCAGTAGCCACTCTTCTTAATAGGCAATATAAGTCAG-----60
QY      235   xGluAsnHisAsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnLeuSe 255
Db      61   -----ACAGAACATATGAAGCTGCTAAACAAGTCGA 92
QY      255   rSerLeuLeuArgGluLeuTyrAlaLysProLeuSer-----GluArgHisValGln 272
Db      93   TGAATATATAAA-----AAAAAGCTCCAATTAGATAGAGAAACACATACCCA 140
QY      272   uSerAspGlyLeuLeuPheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAl 292
Db      141   AAAATTCGGCTTACTACAAAGTTGGCGGTAAATTAACACGAGTATTTCATGGATTAA 200
QY      292   aValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGluLy 312
Db      201   TGTTCATAAAAGAGTCGAGCTGAGTTCGG-----TCAGAAATATAAAGC 248
QY      312   sArgIleAlaArgIleIleProLeuA-gTyrArgSerAsnHisTrpValProAspSerAr 332
Db      249   AAAGTTA-----GACGCAGC 263
QY      332   gProGluGlnProSerProGlnSerThrPro--GluProSerProSerProGlnProAl 351
Db      264   TTTTCAGCAGGTTTATAAAGATACATTACCACAGAACCCAGGAAAAAGGTAGCAGAGC 323

```

```

351 aProAsnProGlnProAlaProSerAsnProIleAspGluLysLeuValLysGluAlaVa 371
324 T-----GAGAAAGAGGTTGAAGAGCTAA 347
371 lArgLysValGlyAspGlyTyrValPheGluAsnGlyValSerArgTyrIleProAl 391
348 GAAAAAAGCCGAGGATCA-----AAAGAAAAGATCTCCGTAACTAC-----CCAAC 395
391 a-----LysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeuAlaLy 408
396 CAATACTTACAAAACGCTTGAACCTTGACATCTGAGTCC-----GATGTGAAGTTAAAAA 452
408 sGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSerSerAspAr 428
453 ACCGAGCTTGAACCTAGTAAAGAGGAGCTAAGGAA-----TCTCGAGACGA 500
428 gGluPheTyrAsnLysAlaTyrAspLeuAlaAlaArgIleHisGlnAspLeuAspAs 448
501 GAAAAAATTAATCAAGCA-----AAAGCGAAAGTTGAGAA 536
448 nLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuGluArgLeuLysAspVa 468
537 TAAAAAAGCTGAGGCT-----ACAAGGTTAAAAACAT 569
468 l**SerAspLysValLysLeuVal**AspIleLeuAlaPheLeuAlaProIleArgHi 488
570 CAAGACATCTGTGAAAA-----GCAGAGAGCTTAACGAGAGC 611
488 sProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIleGlnVa 508
612 AGATGCTAAGTTGCGAAGCTTAATGTAGCGACTTCAGAGCAAGATAAATCAAGAGGCG 671
508 lAlaLys-----LeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspProAr 526
672 GGCAAAACGAGAAGTTCTTGGAGAGCTAGCAACCTGAT----- 711
526 gAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTr 546
711 ----- 711
546 pIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAlaLysGl 566
712 -----GAAAAA-----GAAATGATCGAAGCTTCTCAGATTCATAGCGTAGGTGA 755
566 uLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGl 586
756 AGAAACTCTTACAGCCATCCCTG-----AAACCGAAGAAAAAGGT 797
586 yAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArg----- 604
798 AGCAGAAGCT-----GAGAGAAGGTTGAAGAGCTTAAGAAAAAGCCGAGGATCAAAAA 854
605 -----MetProTyrAsnLeuGlnTyrThr----- 612
855 AGAAGATCGTCTTAACCTACCAACCAATCTTACAAAACGCTTGAATCGAATTCGTGA 914
613 -----ValGluValLysAsnGlySerLeuIleIleProHisTyrAspHisTyrHisAs 630
915 GTCCGATGTGGAAGTTAAAAAGCGGAGCTT----- 945
630 nIleLysPheGluThrPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGl 650
946 -----GAACTAGTAAGAGGAGGAGCTTAAGGAATCTCGAAGAACCA 983
650 uAspLeuAlaThrValLysTyrTyrValGluHisProAsnGluArgProHisSerAs 670
984 GAAAAAATTAAGCAAGTAAAGCGAAGCTTGA-----AGTAA 1022
670 pAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspTh 690
1023 AAAAGCTGAGGCTACAGGCTAGAAAAACATCAAGACAGATCGTAAAAAGCAGAGAGA 1082

```

```

QY 690 rAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluGlu-- 709
Db 1083 AGAAGCTAAACGAGCAGCAGAGAGAGATAAAGTTAAAGAAAAAACCCAGCTGAACACC 1142
QY 710 -----ThrProArgGluGluLysProGlnSerGluLysProGluSerProLy 725
Db 1143 ACAACGAGCGCGCTCTCAACGAGAAACCA-----ACTGAAGAGCTGAGATCCAGC 1199
QY 725 sProArgGluGluProGlu---GluSerProGluGluSerGluGluProGlnValGlu-- 743
Db 1200 TCCAGCTCCAGCTCCAAACACGAGAGATCCA-----GCTGAAAAACCAAGCAGAAAA 1253
QY 744 -----ThrGluLysValGluGlu 749
Db 1254 GCCAGCTGATCAACAAGCTGAAGAA 1278

RESULT 31
US-09-286-981B-31
; Sequence 31, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizenann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cDNA derived
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae
US-09-286-981B-31

Alignment Scores:
Pred. No.: 1,85e-05 Length: 1272
Score: 157.50 Matches: 126
Percent Similarity: 35.89% Conservative: 75
Best Local Similarity: 22.50% Mismatches: 181
Query Match: 3.92% Indels: 178
DB: 4 Gaps: 28

US-09-765-272A-66 (1-763) x US-09-286-981B-31 (1-1272)
QY 216 GlySerArgProSerSer-SerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSe 235
Db 12 GGTAACTATCCCAAGTAGCCACTTCTCTTAATAGGCAATAAAGATCAG----- 60
QY 235 rGluAsnHisAsnLeuThrValThrProTyrHisGlnAsnGlnGlyGluAsnLysSe 255
Db 61 -----ACAGAACATATGAAAGCTGCTTAAACAGTGA 92
QY 255 rSerLeuLeuArgGluLeuTyrAlaLysProLeuSer-----GluArgHisValG1 272
Db 93 TGAATATATATA-----AAAAAGCTCAATAGATAGAGAAAAACATACCCA 140
QY 272 uSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAl 292
Db 141 AAATGTCGGCTTACTCACAAGTTGGCGGTAATTAACCGGAGATTTTGCATGGATTAA 200
QY 292 aValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGluLy 312
Db 201 TGTTTCAAAAAAGAGAGTCGGAAGCTGAGTTGCCG-----TCAGAAATAAAGC 248
QY 312 sArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAspSerAr 332
Db -----

```



Db 249 AAAGTTA-----GACCGAGC 263  
Qy 332 gProGluGlnProSerProGlnSerThrPro---GluProSerProSerProGlnProAl 351  
Db 264 TTTTGAGCAGTTTAAAAAGATACATTACCAACAGAACAGCTAAAGAGGTAGCAGAAGC 323  
Qy 351 aProAsnProGlnProAlaProSerAsnProIleAspGluLeuValValValGluAlaVa 371  
Db 324 T-----GAGAAAGAGGTGGAAGAGCTAA 347  
Qy 371 lArgLysValGlyAspGlyTyrValPheGluGluGlnGlyValSerArgTyrIleProAl 391  
Db 348 GAAAGAGCGGAGGATCAA-----AAAGAAAAGATCTCCGTAACATC---CCAAC 395  
Qy 391 a-----LysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeuAlaLy 408  
Db 396 CAATACTTACAAAACGCTTGAACTTGACATTCGTGAGTCC---GATGTGGAAGTTAANA 452  
Qy 408 sGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSerSerAspAr 428  
Db 453 AGCGAGCTTGAACTAGTAAAGAGAGAGCTAAGGAA-----TCTCGAGACGA 500  
Qy 428 gGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeuAspAs 448  
Db 501 GAAAAAATTATCAAGCA-----AAAGCGAAAGTTGAGAA 536  
Qy 448 nLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLysAspVa 468  
Db 537 TAAAAAGCTGAGGCT-----ACAGGTTTAAAAACAT 569  
Qy 468 l\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeuAlaPheLeuAlaProIleArgHi 488  
Db 570 CAAGACAGATCGTGAANA-----GCAGAAGAGAGCTAAACAGAGAGC 611  
Qy 488 sProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIleGlnVa 508  
Db 612 AGATGCTAAGTTGACGAGGAGTAAATGATGCGACTTCAGACGAGATATAATCAAGGCG 671  
Qy 508 lAlaLys-----LeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspProAr 526  
Db 672 GCGAAAACGAGAAGTCTTGAGAGCTAGCAACCTGAT----- 711  
Qy 526 gAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTr 546  
Db 711 ----- 711  
Qy 546 pIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyrAlaLysGl 566  
Db 712 -----AAAAAA-----GAAAAATGATCGGAAGTCTTCAGATTCCTAGCGTAGGTGA 755  
Qy 566 uLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGl 586  
Db 756 AGAACTCTTACAGCCCATCCCTG-----AAACGAGAAAAAAGGT 797  
Qy 586 yAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArg----- 604  
Db 798 AGCAGAACT---GAGAGAAGGTTGAAGAACTTAAGAAAAAGCCGAGGATCAAAAAGA 854  
Qy 605 -----MetProTyrAsnLeuGlnTyrThr----- 612  
Db 855 AGAAGATCGTGTACTACTACCAACCAATACTTACAAAACGCTTGAACTTGAATTCCTGA 914  
Qy 613 -----ValGluValLysAsnGlySerLeuIleIleProHisTyrAspHisThrHisAs 630  
Db 915 GTCCGATGTGGAAGTTAAAAAGCGGAGCTT----- 945  
Qy 630 nIleLysPheGluTyrPheAspGluGlyTyrGluAlaProLysGlyTyrThrLeuGl 650  
Db 946 -----GAAGTAGTAAAGAGGAGGAGCTAGGATCTCGAAACGA 983  
Qy 650 uAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGluArgProHisSerAs 670  
Db 984 GGAATAAATTAAGCAAGTAAAGCAAGGAAAGTTGAG-----AGTAA 1022

Qy 670 pAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspTh 690  
Db 1023 AAAAGCTGAGCTACAGGCTAGAAAAACATCAACACAGATCGTAAAAAAGCAGAGAAGA 1082  
Qy 690 xAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluGlu-- 709  
Db 1083 AGAAGCTAAAGCAAGCAGCAGCAAGAAAGATAAAGTTAAAGAAAAAACAGCTGAACAACC 1142  
Qy 710 -----ThrProArgGluGluLysProGlnSerGluLysProGluSerProLy 725  
Db 1143 ACAACAGCGCGCTCTCTCAACAGAAAAACCA---ACTGAGAGCTTGAGATCCAGC 1199  
Qy 725 sProThrGluGluProGlu---GluSerProGluGluSerGluGluProGlnValGlu 743  
Db 1200 TCCAGCTCCAGCTCCAAAAACAGAGAAATCCA-----GCTGAAAAAACCAAAAGCAGAA 1251  
RESULT 32  
US-09-286-981B-30  
; Sequence 30, Application US/09286981B  
; Patent No. 6503511  
; GENERAL INFORMATION:  
; APPLICANT: Wizenmann, Theresa M.  
; APPLICANT: Koenig, Scott  
; APPLICANT: Johnson, Leslie S  
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines  
; FILE REFERENCE: 469201-396  
; CURRENT APPLICATION NUMBER: US/09/286,981B  
; CURRENT FILING DATE: 1999-04-06  
; PRIOR APPLICATION NUMBER: US 60/085,743  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 30  
; LENGTH: 1276  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:cDNA derived  
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae  
US-09-286-981B-30

Alignment Scores:  
Pred. No.: 2,32e-05 Length: 1276  
Score: 156.50 Matches: 126  
Percent Similarity: 35.89% Conservative: 75  
Best Local Similarity: 22.50% Mismatches: 181  
Query Match: 3.89% Indels: 178  
DB: 4 Gaps: 28

US-09-765-272A-66 (1-763) x US-09-286-981B-30 (1-1276)

Qy 216 GlySerArgProSerSer-SerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSe 235  
Db 12 GGTAACCTACCAAGTAGCCACTTCTCTTAATAGCAATAATAAAGTCAG----- 60  
Qy 235 xGluAsnHisAsnLeuThrValThrProTyrTyrHisGlnAsnGlnGluAsnIleSe 255  
Db 61 -----ACAGAACATATGAAGCTGCTAAACCAAGTCGA 92  
Qy 255 rSerLeuLeuArgGluLeuTyrAlaLysProLeuSer-----GluArgHisValGl 272  
Db 93 TGAATATATAAAA-----AAAAAGCTTCAATTAGATAGAGAAACATACCCA 140  
Qy 272 uSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAl 292  
Db 141 AAATGTCGGCTTACTCACAAGTTGGCGTAATATAAACCGAGATTTGCGATGATTAAAG 200  
Qy 292 aValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGly 312  
Db 201 TGTTCAAAAAAGAGTCGGAAGCTGAGTTGCCG-----TCAGAAAAATAAAGC 248  
Qy 312 sArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAspSerAr 332

Db 249 AAGTTA-----GAGCAGC 263  
QY 332 gProGluGlnProSerProGlnSerThrPro---GluProSerProSerProGlnProAl 351  
Db 264 TTTTGGAGCTGTTTAAAAAGATACATACCACAGACACCAAGGAAAAAGGTAGCAGAAC 323  
QY 351 aProAsnProGlnProAlaProSerAsnProAlaAspGluLysLeuValLysGluAlaVa 371  
Db 324 T-----GAGNAGAGGTTGAGNAGCTAA 347  
QY 371 lArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSerArgTyrIleProAl 391  
Db 348 GAAAAAGCGCGAGGATCAA-----AAAGAAAAAGATCTCGTAATAC---CCAAC 395  
QY 391 a-----LysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeuAlaLy 408  
Db 396 CAATATTCAAAACCGCTTGAATGATGCTGAGTCC---GATGTGAAGTTAAAAA 452  
QY 408 sGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSerSerAspAr 428  
Db 453 AGCGGAGCTTGAAGTAAAGAGGAGCTTAGGAA-----TCTCAGACGA 500  
QY 428 gGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeuAspAs 448  
Db 501 GAAAAAATTAATCAAGCA-----AAAGCGAAAGTTGAGAA 536  
QY 448 nLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuGluArgLeuLysAspVa 468  
Db 537 TAAAAAAGCTGAGGCT-----ACAAGGTTAAAAAATCAT 569  
QY 468 l\*\*\*SerAspLysValLysLeuVal\*\*\*AspLleLeuAlaPheLeuAlaProIleArgHi 488  
Db 570 CAAGCAGATCGTGAAGAA-----GCAGAGAGCTAAACGAGAGC 611  
QY 488 sProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIleGlnVa 508  
Db 612 AGATGCTAAGTTCGAGAGAGCTAATGAGCAGCTTCAGAGCAAGATAAATCAAGAGCG 671  
QY 508 lAlaLys-----LeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAspProAr 526  
Db 672 GGCAGAACGAGAGGTTTTCGAGAGCTAGCAACACCTGAT----- 711  
QY 526 gAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTr 546  
Db 711 ----- 711  
QY 546 pIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAlaLysGl 566  
Db 712 -----GAAATGATGCGAGCTTCAGATTCAGATTCAGCTAGGTGA 755  
QY 566 uLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGl 586  
Db 756 AGAACTCTTACAAAGCCATCCCTG-----AAACAGAGAAAAAGGT 797  
QY 586 yAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArg----- 604  
Db 798 AGCAGAGCT---GAGAGAGAGGTTGAAGAGCTTAAGAAAAAGCCGAGGATCAAAAAA 854  
QY 605 -----MetProTyrAsnLeuGlnTyrThr----- 612  
Db 855 AGAAGATGCTGCTAATACCAACCAATACCTTCAAAACGCTTGAACTTGAATTCGTGA 914  
QY 613 -----ValGluValLysAsnGlySerLeuIleIleProHisTyrAspHisAs 630  
Db 915 GTCGAGTGTGGAAGTTAAAAAGCGGAGCTT----- 945  
QY 630 nIleLysPheGluTrpPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGl 650  
Db 946 -----GAACTAGTAAAGAGAGGAGCTAAGGAATCTCGAAACGA 983  
QY 650 uAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGluArgProHisSerAs 670  
Db ----- 711

Db 984 GMAAAATTAAGCAAGTAAAGACGAAAGTTTCAG-----AGTAA 1022  
QY 670 pAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaaspTh 690  
Db 1023 AAAAGCTGAGGCTACAAGGCTAGAAAACATCAAGACAGATCGTAAAAAAGCAGAGAAGA 1082  
QY 690 rAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluGlu-- 709  
Db 1083 AGAGCTAAACGAGAGCAGCAGAGAGAGATAAAGTTAAAGAAAAAACCCAGCTGAACACC 1142  
QY 710 -----ThrProArgGluGluLysProGlnSerGluLysProGluSerProLy 725  
Db 1143 ACAACGAGCGCGGCTCTCAACGAGAAAAACCA---ACTGAAGCGCTGAGATCCAGC 1199  
QY 725 sProThrGluGluProGlu---GluSerProGluGluSerGluGluProGlnValGlu 743  
Db 1200 TCCAGCTCAGCTCCAAAAACGAGAGATCCA-----GCTGAAAAACCAAGACGAGAA 1251

## RESULT 33

US-08-961-527-134  
; Sequence 134, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 134:

SEQUENCE CHARACTERISTICS:

LENGTH: 12665 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-134

## Alignment Scores:

Pred. NO.: 0.000908 Length: 12665  
Score: 156.50 Matches: 130  
Percent Similarity: 35.70% Conservative: 76  
Best Local Similarity: 22.53% Mismatches: 213  
Query Match: 3.89% Indels: 158  
DB: 4 Gaps: 28

US-09-765-272A-66 (1-763) x US-08-961-527-134 (1-12665)

QY 214 LysGlnGlySerArgProSerSer-SerSerSerTyrAsnAlaAsnProAlaGlnProAr 233

Db 687 GAACGAGGAGCTACCCCAAGTACCACCTTCTTCTTAATAGGCGCAATGAAGATCAGGAGCA 746

233 gLeuSerGluAsnHisAsnLeu-----ThrValThrProThrTyrHisGlnAsnGlnG1 251  
747 ACAAGGAGAAACAACCTAAAAAATCGATTTCAGAACGAGTAAGCGAAGAGAGGTGCA 806  
251 yGluAsnLeuSerSerLeuLeuAArgGluLeuTyrAlaLysProLeuSerGluAArgHisVa 271  
807 GGAATATGTAATAAATAATAGGGGTGAGAGCTATGCAAAAATCAATAAAACGCGACATAC 866  
271 lGluSerAspGlyLeuLeuPheAspProAlaGlnIleThrSerArgThrAlaAArgGlyVa 291  
867 AATTACTAGTCTAGTCTAAGCAGAGTTCAACCAACATTAAAG----- 906  
291 lAlaValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuG1 311  
907 -----AACGAGTATTGTAATAAATAGTTGAATCAACCTCAGAAAGCCA 950  
311 uLysAArgIleAlaArgIleLeuProLeuAArgTyrArgSer-----As 325  
951 A-----CTACAGATACTGATGATGAGAGTGCATCAAAAGTAGATGAAGCTGTGTC 1001  
325 nHisTrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSe 345  
1002 TAAGTTTGAAGAGCACTCATCTTCTGTCAGTTCAGACTCTTCCACT-----AAACCCGA 1058  
345 rProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGlu-- 364  
1059 AGCTTCAGATACAGCGAAGCAACAAAGCCAGACAGCAAGAGAAAGGTAGCAGAGC 1118  
365 ---LysLeuValLysGluAlaValAArgLysValGlyAspGlyTyrValPheGluGluAs 383  
1119 TAAGAAAGAGGTGAAGAGCTGAGAAAGCAAGCCAGAGTCAAA-----AAAGAGA 1169  
383 nGlyValSerArgTyr-----IleProAlaLysAspLeuSerAlaGluThrAla--Al 400  
1170 AGATCGTCGTAACTACCAACCATTAATTACAAACCGCTTGAATTTGCTGAGTC 1229  
400 aGlyIleAspSerLysLeuAlaLysGlnGlnSerLysLeuGlyAlaLysLY 420  
1230 CGATGTGGAAGTTAAAAAGCGGAGCTTGAA-----CTAGTAAAGTGAA 1274  
420 sThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaAr 440  
1275 AGCTAACGAACCTCGAGACAGCAAAATTT----- 1305  
440 gIleHisGlnAspLeuLeuAspAsnLysGlyAArgGlnValAspPheGluAlaLeuAspAs 460  
1306 -----AAGCAAGCAAGAGCGGAGTGAAGTGAAGAGCAAGAGAA----- 1395  
460 nLeuLeuGluAArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLe 480  
1341 TGAGGCTACAGGTAAATAAATAATCAAGACAGATCGTGAAGACAGCAAGAGAA----- 1395  
480 uAlaPheLeuAlaProIleArg--HisProGluAArgLeuGlyLysProAsnAlaGlnI1 499  
1396 -----GCTAAACGAGAGCAGATGCTAAAGAGCAAGAGTAAACCAAGGGGGCGG-- 1443  
499 eThrTyrThrAspAspGluLeuGlnValAlaLys-----LeuAlaGlyLysTyrThrTh 517  
1444 -----GCAAAACGAGGAGTCTCTCGAGAGCTAGCAAC 1475  
517 rGluAspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrVa 537  
1476 ACCTGAT----- 1482  
537 lThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluAr 557  
1483 -----AAAAA-----GAAATGATGC 1499  
557 gAlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisG1 577  
1500 GAAGTCTTCAGATTCAGGTAGGTGAAGAACTCTTCCAAGGCCATCCCTG----- 1551

577 nAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnAArgValLysAlaAl 597  
1552 -----AAACCAAGAAAAAGGCTAGCAGAAGCT---GAGAAGAAAGGTTGAAGAGC 1598  
597 aLysLysValProLeuAspArg-----MetProTyrAsnLeuG1 610  
1599 TAAGAAAAAAGCGAGGATCAAAAAGAGAAGAGATCGCGTAACCTACCAACCAATACTTA 1658  
610 nTyrThr-----ValGluValLysAsnGlySerLeuI1 621  
1659 CAAAACGCTTGAACTGAATTTGCTCAGTCCTGATGCGAAGTTAAAAAACGGGAGCTT-- 1716  
621 eIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeuTy 641  
1717 -----GAACCTAGTAAA 1727  
641 rGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValG1 661  
1728 AGAGGAAGCTAAGGAACCTCGAAACGAGGAAAAAGTTAAAGCAAGCAAAAGCGGAAGTTGA 1787  
661 uHisProAsnGluAArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValG1 681  
1788 G-----AGTAAAAAGCTGAGGCTACAGGTTAGAAAAATCAA 1826  
681 nArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysPr 701  
1827 GACAGATCGTAAA---AAAGCAGAGAAGCAAGCTAAACCGAAAAGCAGCAGAGAGATAA 1883  
701 oGlnThrGluLysProGluGluGluThrProAArgGluGluLysProGlnSerGluLysPr 721  
1884 AGTTAAAGAAAAACCAAGCTGAACCAACCAAGCGCGCGCTCCAAAAGCAGAAAAACC 1943  
721 o---GluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPr 740  
1944 AGCTCCAGCTCAAAACCA---GAGAATCCAGCTGAACCAACCA---AAAGCAGAAAAACC 1997  
740 oGlnValGluThrGluLysValGluLysLeuAArgGluAlaGluAsp 756  
1998 A---GCTGATCAACAGCTGAAGAGACTATGCTCGTAGATCAGAGAA 2043

## RESULT 34

US-08-961-527-83/c  
; Sequence 83, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504

52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069

; TITLE OF INVENTION: Nucl

APPLICANT: Fraser et al.

; APPLICANT: Fraser et al.



616 ----- 616  
485865 AAATTGATTCACCACTGTTGAACACCTGAAATTTAGTAGTGAACCACTCAA 485806  
617 -----AsnGlySerLeuIleProHisTyrAspHisTyrHisAsnIle 631  
485805 CAACATGTTGAATCAGAACTGTTTGTATGATGAACCTTAATATATGAC----- 485761  
632 LysPheGluTTPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAsp 651  
485760 -----TTTGATGAACCTTAACCTATGATTTGTATGATCAACCTTAGTTATGATTCAGAT 485713  
652 LeuLeu---AlaThrValLysTyrTyrValGluHisProAsnGluArgProHisSerAsp 670  
485712 CTACACCATCAGAACCTCAATATGATGTAGATGAACCTTAAT----- 485671  
671 AsnGlyPheGlyAsnAlaSerAspHisValGln---ArgAsnLysAsnGlyGlnAlaAsp 699  
485670 TATGACTTTGTATGAACCAAAATATGAAATTTGAATCAAAACCATCAGAGCCACAATTTGAA 485611  
690 ThrAsnGlnThrGluLysProSerGlu----- 698  
485610 COTCAAGTAGAACCAACCTCGAGAACGAGTCTTTTGAACCATCTCAGAACGCTAAATTT 485551  
699 -----GluLysProGlnThrGluLysProGluGlu-----Glu 709  
485550 GATTCACCACTGATCGTTCAAGATTCACACCTGAACTCTCTTGAAGAGTTCAA 485491  
710 ThrProArgGluLysProGlnSerGluLysProGluSerProLysProThrGluGlu 729  
485490 ACTCAACAGAAATCAACCTGTAGAATCTCAACCAAGAGTACTTTTGTATGATCTGTTCAA 485431  
730 ProGluGluSerProGluGluSer-----GluGluPro 740  
485430 COTGAACAACTCTCAGAACGCTAAATTTGATTCACAGTTGAAACTATTCAGAACCT 485371  
741 GlnValGluThrGlu 745  
485370 CAAGTTCTAGTGAA 485356

RESULT 36  
JS-08-923-992A-5  
Sequence 5, Application US/08923992A  
Patent No. 6280728  
GENERAL INFORMATION:  
APPLICANT: Tai, Joseph Y.  
APPLICANT: Blake, Milan S.  
TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B  
TITLE OF INVENTION: Streptococcal Beta Antigens  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/923,992A  
FILING DATE: 05-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/024,707  
FILING DATE: 06-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 1438.0140001/RWE  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3384 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..3384  
US-08-923-992A-5

Alignment Scores:  
Pred. No.: 0.000369 Length: 3384  
Score: 151.00 Matches: 164  
Percent Similarity: 35.14% Conservative: 128  
Best Local Similarity: 19.74% Mismatches: 317  
Query Match: 3.76% Indels: 222  
DB: 3 Gaps: 44

US-09-765-272A-66 (1-763) x US-08-923-992A-5 (1-3384)

QY 8 HisGlnAlaGlyGlnValLysLys-----GluSerAsnArgValSerTyrIleAspGly 25  
Db 490 CATCAAAAAAGTCCAGTTGAAAAAATGCGAGGCAAAAGGGAATCACAATGAAGATAAA 549  
QY 26 Asp-----GlnAlaGlyGlnLysAlaGluAsnLeu 35  
Db 550 GATTCATGCTGAAAAAATCGAAGATATTCGTAACCAAGCTCAACAGCAGATAAAAA 609  
QY 36 ThrProAspGluValSerLysArgGlu----- 44  
Db 610 GAAGATGCCGAAGTAAAGTTGCTGAGAACCTAGGTAAACTCTTTAGTTCAACTAAAGCT 669  
QY 45 GlyIleAsnAlaGlu-----Gln\*\*\*ValIleLysIleThrAspGlnGlyTyrValThr 62  
Db 670 GGTCTGGATCAAGAAATTCATGAGCATGTGTGAAGAAAGAACGAGTAGTGAGGAAATACT 729  
QY 63 SerHisGlyAspHisTyrHisTyrTyrAsn-----GlyLys 74  
Db 730 CAGAAAGTTGATGAA---CACTATGCTTAATAGCCCTTCAGAACCTGCTCAAAAATCTCTT 786  
QY 75 ValProTyrAspAlaIleIleSerGluGluLeuMetLysAspProAsnTyrGlnLeu 94  
Db 787 GAAGAACTAGATAAGGCAACTACCAATGAACAGCTACACAAGTTAAAAATCAATCTTTA 846  
QY 95 LysAspSerAspIleValAsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLys 114  
Db 847 GAAACCGCTCAAAAGCTCAAAAGAAATGCAA---CCTCTTATCAAGAAACCAAGTAATGAAA 903  
QY 115 TyrTyrValTyrLysLysAspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIle 134  
Db 904 TTGTATAAGGCTATG-----AGTGAGAGCTTG 930  
QY 135 LysArgGlnLysGlnGluArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAla 154  
Db 931 GAGCAGTTGAGAGGAAGTAATAAACATAATTCGAGAGCTAATTTAGAGATTGTTGGTGG 990  
QY 155 AlaAlaArgAlaGlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAsp 174  
Db 991 AAATCTAAA-----GAA 1002  
QY 175 IleIleGluAspThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIle 194  
Db 1003 ATCGTAAGAGAA-----TACGAGAGAAA 1026  
QY 195 ProLysAsnGluLeuSerAlaSerGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLys 214  
Db 1027 CTTAATCAATCTAAAAATCTTCGAAATTTAAAGCACTAGAAAGAGGAGCTCATTCGAG 1086





```

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PH340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 145:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-145

Alignment Scores:
Pred. No.: 0.00233 Length: 10711
Score: 151.00 Matches: 170
Percent Similarity: 32.92% Conservative: 119
Best Local Similarity: 19.36% Mismatches: 340
Query Match: 3.76% Indels: 249
DB: 42 Gaps: 42

JS-09-765-272A-66 (1-763) x US-08-961-527-145 (1-10711)
QY 5 LeuGlyArgHisGlnAlaGlyGlnValLysGluSerAsnArgValSerTyrIleAsp 24
DB 444 TTGGATGCGCTTATAAAGGGGAATCAAA-----GAT 476
QY 25 GlyAspGlnAlaGlyGlnLysAla-----GluAsnLeuThrProAspGluValSerLys 42
DB 477 TTCTCAAGTAGGAATTAAAGGCTCTGGATGAACAGACAGCTTCAGTACACTTTGAACAAA 536
QY 43 ArgGluGlyLe-----AsnAlaGluGln**ValIleLysIleThrAspGlnGly 59
DB 537 CCAGAAAGCTTCGGAATCTTAAGACAAACCATGGTGCTGGCCAGTAAAGAGAG 596
QY 60 TyrValThrSerHisGlyAspHisTyrHis-----TyrTyr 71
DB 597 TTTTGAATTCAAAGAGAGATGATTTGGCCAAAGCTACGGATCCAAAGTAGTCTCTGTAT 656
QY 72 AsnGlyValProTyrAspAlaIleIleSerGluGluLeu-----MetLysAsp 89
DB 657 AACGGTCTTATTGTTGAAATCCATTTGACCAAAATCCTGTGTGAATTTGCGAAAAAT 716
QY 90 ProAsnTyrGlnLeuLysAspSerAspIleValAsnGluIleLysGlyGlyValIle 109
DB 717 CCGAATCTCTGGATGAAGCAAT-----GTGCATGTTGAC 752
QY 110 LysValAsnGlyLysTyrTyrValTyrValTyrLysAspAlaAlaHisAspAsnIleArg 129
DB 753 AAAGTTAAATTCATTCTGGATGGTCAAGATACCAGCAACCTGCAGAAAACTTTAAA 812
QY 130 -----ThrLysGluGluIleLysArgGlnLysGlnGluArgSerHisAsnHis 145
DB 813 GATGGTAGCCTTACAGCAGCTCTCTCTATCCAAAGTCCAAAGTTCGCGAAGACTTGAG 872
QY 146 AsnSerArgAlaAspAsnAlaValAlaAlaAlaAlaGlnGlyArgTyrThrAsp 165
DB 873 ACAGATGATGAAGCAATATTGTC-----TATACCTCAACAA 908
QY 166 AspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleVal 185
DB 909 GACTCTATTACGTATCTAGTTGGTACAAATATTGACCGTCAGTCTCTATAAATACATCT 968
QY 186 ProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAla 205
DB 969 AAGACCAAGCAGCAACAAAGAGGATGACTTAAAGGCTCTCTTAAACAAAGATTCGGT 1028
QY 206 AlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerTyr 225
DB 10711

```

```

DB 1029 CAGGCTATTGCGCTTT-----GGATTGACCGTACACGCTATGCTCTCAGTTG 1076
QY 226 AsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThr 245
DB 1077 AATGGACAACTGGAGCAAGTAATCTTG-----CGTAATCTCTTTGTGGCCACCAACA 1130
QY 246 TyrHisGlnAsnGlnGlyAsnIleSerSerLeuLeuArgGlu-----LeuTyr 262
DB 1131 TTTGTTCAAGCAGATGTTAAACCTTTGGCGATATGGTCAAGAGAAATTTGGTCACTTAT 1190
QY 263 AlaLysProLeuSerGluArgHisVal-----GluSerAspGlyLeuIlePheAspPro 280
DB 1191 GGGGATGAATGGAAGGATGTTAATCTTGAGATTTCTCAGGATGCTT-----TACAATCCA 1247
QY 281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300
DB 1248 GAAAAAGCCAAAGGCTGAATTTGCTAAAGCTAATCAGCCTTAAACGACGAGGAGTCCAA 1307
QY 301 Phe-----IleProTyrGluGlnMetSerGluLeuGluLysArgIleAla 315
DB 1308 TTCCCAATTCATTGGATATGCCAGTTGACCAACAGCA-----ACTACAAAAGTTCAG 1361
QY 316 ArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGln 335
DB 1362 CGCGTC----- 1367
QY 336 ProSerProGlnSerThrProGluProSerProSerProGlnProAlaProAsnProGln 355
DB 1368 -----CAATCTATGAACCAATCCTTGGAGCAACTTTAGGAGCTGATATATGTC--- 1415
QY 356 ProAlaProSerAsnProIleAsp---GluLysLeuValLysGluAlaValArgLysVal 374
DB 1416 -----ATTATTGATATTCAACAACTACAAAAGACGAAGTAACAATATT 1460
QY 375 -----GlyAspGlyTyrValPheGluGluAsn----- 383
DB 1461 ACATATTTTGTGTAATAATGCTGCTGCGAAGACTGGGATTTTATCAGATAATGTCGGTTGG 1520
QY 384 -----GlyValSerArgTyrIle-----ProAlaLysAspLeu 394
DB 1521 GGTCCAGACTTTTGGCGATCCATCACTACCTTGATATTATCAAACTCTCTGTAGGAGAA 1580
QY 395 SerAlaGluThrAlaAlaGlyIleAspSer----- 404
DB 1581 AGTACTAAACATATTTAGGGTTTCACTCAGCGGAGATAATGCTAGCTGCTAAAAAGTA 1640
QY 405 LysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuPro 424
DB 1641 GGTCTATATGACTACGAAAAATTTGGTTACTGAGGCTGGTGATGAGACTACAGATGTTGCT 1700
QY 425 SerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAsp 444
DB 1701 -----AAAGCTATGATAAATACGCTGACGACCCCAAGCTTGG 1736
QY 445 LeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArg 464
DB 1737 TTGACAGATAGT----- 1748
QY 465 LeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeuAlaPheLeuAla 484
DB 1749 -----GCTTTGATTATT 1760
QY 485 ProIleArgHisProGluArgLeuGlyLysPro-----AsnAlaGlnIleThrTyrThr 502
DB 1761 CCAACT-----ACATCTCGTACAGGGCGCTCAATCTTGTCTCAAGATGCTACCATTTACA 1814
QY 503 AspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAsp----- 519
DB 1815 -----ATACCATTTGCAATTTGTGAGAAATAAAGGTACAAAGTGAACCCAGTCTTGTAT 1865
QY 520 GlyTyrIle---PheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
DB 1866 AAATACCTTGGAACTTCAAGACAAAGGCGACTGCTAGATGAA----- 1907

```

```
QY 539 ProHisMetThrHisSerHisTriPileLysLysAspSerLeuSerGluAlaGluArgAla 558
Db 1908 TACCAAAAAGCTCAGGAAAAATGGATGAAGAAAA-----GAAGAGTCTAATAAAAAAG 1961
QY 559 AlaAlaGlnAlaTyraLys-----GluLysGlyLeu 569
Db 1962 GCTCAAGAAGATCTCGCAAAACATGTGAAATAACTGTTGCAAAATATAAGAAAGATTTA 2021
QY 570 ThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGlyAlaGluAla 589
Db 2022 GTATTTCCCTTG-----AATGCTGAATCCCTTTTATCAATTTGTAAGAAAGATCTTAAA 2075
QY 590 IleTyraAsnArgValLysAlaAlaLysLysValProLeuAspArgMetProTyraAsnLeu 609
Db 2076 ATGTACGACCCCAAGAGTTGGAGCCCTCTTTTGTGCAGATAGAGAAATTTTGTGTTAA 2135
QY 610 GlnTyraThrValGluValLysAsnGly-----SerLeuIleIle 622
Db 2136 TTTTACTTGTTCCTAATGCTTTCTCAGCTATTATTGTTATATTATAAAGTATAATTATT 2195
QY 623 ProHis-TyraSp-----HisTyraHisAsnIleLysPhe----- 633
Db 2196 TTTTATTATCAGAGTTAAGCATTCGACTTTTCAGAGGAGGAGTATTTTAAAAAGAAA 2255
QY 634 -----GluTrpPheAspGluGlyLeuTyraG1 642
Db 2256 ATGTAAACGTTTGTCTCAAAATGAAGGATTAGAGTTTATGATTAAGGATTTATTGTA 2315
QY 642 uAlaProLysGlyTyraThrLeuGlu-----AspLeuLeuAlaTh 655
Db 2316 AAAACGTTGTAAATATAGTATTTCGAAATTTTCATTAGTGTGCTTCTGTTATGATTGG 2375
QY 655 rValLysTyraThrValGluHisProAsnGluArgProHisSerAspAsnGly----- 672
Db 2376 AGTGCATTCTTTGGGCAAGTCCGGTCTTTCAGATAGCGTGCAGTCTGTTCCACGGC 2435
QY 673 -----PheGlyAsnAlaSerAspHisValGlnArgAs 683
Db 2436 GAACTTACCAGCTGATTAGCTACTGCTCTTGCACACGAAAGAGATGATGGCGGTGA 2495
QY 683 n-----LysAsnGly-----GlnAlaAspThrAsnGlnThrGluLysProSe 697
Db 2496 TTTTGAAGCGCTAAGTGGGAGAGACCAAGTTCTCCAGAAAGTTACAGATGGACCTAA 2555
QY 697 rGluLysPro-----GlnThrGluLysProGluGluThrProArgG1 713
Db 2556 GACAGAAGAAGAACTATTAGCACTTGTAAGAAAGAAACCGGCTGAAGAAAGAAACCAAGA 2615
QY 713 uGluLysProGlnSerGluLysProGluSerProLysProThrGluGluProGluGluSe 733
Db 2616 GGATTAACCTGCAGCTGCTAAACCTGAACACCTTAAG---ACGGTAACCCCTGAATGGCA 2672
QY 733 rProGluGluSerGluGluProGlnValGluThrGluLysValGluGluLys 750
Db 2673 AACGGTAGCGAATAAGAGCAACAGGGAACAGTCACTATCCGAGAGAAGAAA 2724
```

## RESULT 38

```
US-09-286-981B-26
; Sequence 26, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizenann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 469201-396
; CURRENT APPLICATION NUMBER: US/09/286,981B
; PRIORITY FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 26
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cDNA derived
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae
US-09-286-981B-26
```

```
Alignment Scores:
Pred. No.: 9,35e-05 Length: 1338
Score: 150.50 Matches: 126
Percent Similarity: 34.92% Conservative: 72
Best Local Similarity: 22.22% Mismatches: 213
Query Match: 3.74% Indels: 156
DB: 4 Gaps: 26

US-09-765-272A-66 (1-763) x US-09-286-981B-26 (1-1338)
QY 214 LysGlnGlySerArgProSerSer-SerSerSerTyraAsnAlaAsnProAlaGlnProAr 233
Db 6 GAACGAGGAGGAGTACCAGTAGTACCACCTTCTTAATAGGCAATGAAAGTCAGGCAGA 65
QY 233 gLeuSerGluAsnHisAsnLeu-----ThrValThrProThrTyraHisGlnAsnGlnG1 251
Db 66 ACAGGAGAGAACCAACCTAAAAAACTCGAATTCAGAACGAGATAAGCGCAAGGAAGAGGTGCA 125
QY 251 yGluAsnIleSerSerLeuLeuArgGluLeuTyraLysProLeuSerGluArgHisVa 271
Db 126 GGAATGATGAAAAAATAGTGGTGAGAGCTATGCAAAATCAATCAATCAATAAAGGACATAC 185
QY 271 lGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyVa 291
Db 186 AATTACTGTAGCTCTAGTTAACGAGTTGACACATTAG----- 225
QY 291 lAlaValProHisGlyAsnHisTyraHisPheLeuProTyraGluGlnMetSerGluLeuG1 311
Db 226 -----AACGAGTATTGTAATAAATAGTTGTAATCAACCTCAGAAAGCCA 269
QY 311 uLysArgIleAlaArgIleLeuProLeuArgTyraArgSer-----As 325
Db 270 A-----CTACAGATACTGATGATGGAGAGTCCGATCAAAAGTAGATCAAGCTGTGTC 320
QY 325 nHisTrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSe 345
Db 321 TAAATTGGAAGAGGACTCATCTCTTCGTCAGCTCTCCACT---AAACCGGA 377
QY 345 rProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProLysAspGlu-- 364
Db 378 AGCTTCAGATACAGCGAAGCCAAACAAAGCCGACAGAACCAAGGAGAGAGGTAGCAGAAAGC 437
QY 365 ----LysLeuValLysGluAlaValArgLysValGlyAspGlyTyraValPheGluGluAs 383
Db 438 TAAAGAAGAGGTTGAAGAAGCTGAGAAAGAAAGCCCAAGGATCAA-----AAAGAGA 488
QY 383 nGlyValSerArgTyra-----IleProAlaLysAspLeuSerAlaGluThrAla---Al 400
Db 489 AGATCGTCGTAACCTACCCCAACCATTTACTTACAAACGCTTGAACCTTGAATTCGTGAGTC 548
QY 400 aGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLy 420
Db 549 CGATGTGGAAGTTAAAAAGCGGAGCTTGAA-----CTAGTAAAGTGAA 593
QY 420 sThrAspLeuProSerSerAspArgGluPheTyraAsnLysAlaTyraAspLeuLeuAlaAr 440
Db 594 AGCTAACGAACCTCGAGACGACGCAAAAAAT----- 624
QY 440 gIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAs 460
Db 625 -----AAGCAAGCAGAGCGGAAGTTCAGAGTCAACCAAGC 659
QY 460 nLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLe 480
```

660 TGAGCTACAGGTTAAAAAATCAAGACAGATCGTGAAGAGCAGAGAGAA----- 714  
480 uAlaPheLeuAlaProIleArg---HisProGluArgLeuGlyLysProAsnAlaGlnIle 499  
715 -----GCTAAACGAGAGCAGATGCTTAAGAGCAGAGTAAACCAAGGGGGG-- 762  
499 eThrTyrThrAspAspGluIleGlnValAlaLys-----LeuAlaGlyLysTyrThr 517  
763 -----GCAAAACGAGGAGTTCCTGGAGAGCTGCAAC 794  
517 rGluAspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrVa 537  
795 ACCTGAT----- 801  
537 lThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluAr 557  
802 -----AAAAA-----GAAATGATGC 818  
557 gAlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisG1 577  
819 GAAGCTTCAGATTCTAGCTAGGTGAGAACTCTTCCAGCCCATCCTG----- 870  
577 nAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAl 597  
871 -----AAACCCAGAAAAAAGGTAGCAGAGCT---GAGAAGAAGGTTCAGAGAAGC 917  
597 aLysLysValProLeuAspArg-----MetProTyrAsnLeuG1 610  
918 TAAGAAAAAAGCCGAGGATCAAAAAAGAGAGATCCCGTAACTACCCAACTACTTA 977  
610 nTyrThr-----ValGluValLysAsnGlySerLeuI1 621  
978 CAAACGCTTGAACTGAAATTCCTGAGTCCGATGTTGAAGTTAAAGACCGGAGCTT-- 1035  
621 eileProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeuTy 641  
1036 -----GAACTAGTATAA 1046  
641 rGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValG1 661  
1047 AGAGGAGCTAAGGAACCTCGAAACCGAGGAAAGTTAAGCAAGCAAAACCGGAAGTTGA 1106  
661 uHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValG1 681  
1107 G-----AGTAAAAAGCTGAGGCTACAGGTGAGTAAAGAAATCAA 1145  
681 nArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysPr 701  
1146 GACAGATCGTAAA---AAAGCAGAGAGAGCTAAACGAAAGAGCAGCAGAGAGATAA 1202  
701 oGlnThrGluLysProGluGluThrProArgGluGluLysProGlnSerGluLysPr 721  
1203 AGTTAAAGAAAAAACCAAGCTGAACAACACCAACAGCGCGCGCTCCAAAGCAGAAAAACC 1262  
721 o---GluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPr 740  
1263 AGCTCAGCTCCAAAACCA---GAGAAATCCAGCTGAACAACCAAGAGAGAGAGAGAGAG 1319  
740 oGlnValGluThrGluLys 746  
1320 TGATCAACAGCTGAAGAA 1338

RESULT 39  
US-09-535-008-66  
Sequence 66, Application US/09535008  
Patent No. 645629  
GENERAL INFORMATION:  
APPLICANT: Wong, Alexander K.C.  
APPLICANT: Tavtigian, Sean V.  
APPLICANT: Teng, David H.-F.  
TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE  
TITLE OF INVENTION: AND OTHER CANCER TYPES  
FILE REFERENCE: 2318-259

CURRENT APPLICATION NUMBER: US/09/535,008  
CURRENT FILING DATE: 2000-03-23  
EARLIER APPLICATION NUMBER: U.S. 60/125,806  
EARLIER FILING DATE: 1999-03-23  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 66  
LENGTH: 5468  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (75)..(5012)  
FEATURE:  
NAME/KEY: conflict  
LOCATION: (1780)  
OTHER INFORMATION: GenBank Accession No. 645629 U29175 shows a C at this  
OTHER INFORMATION: position (position 1784 in GenBank) rather than  
OTHER INFORMATION: the G shown here.  
FEATURE:  
NAME/KEY: allele  
LOCATION: (26)  
OTHER INFORMATION: Polymorphism of either T or C in this noncoding  
OTHER INFORMATION: region.  
FEATURE:  
NAME/KEY: allele  
LOCATION: (1583)  
OTHER INFORMATION: Polymorphism of A or G resulting in a silent  
OTHER INFORMATION: mutation.  
FEATURE:  
NAME/KEY: allele  
LOCATION: (1598)  
OTHER INFORMATION: Polymorphism of T or C resulting in a silent  
OTHER INFORMATION: mutation.  
FEATURE:  
NAME/KEY: allele  
LOCATION: (1892)  
OTHER INFORMATION: Polymorphism of A or G resulting in a silent  
OTHER INFORMATION: mutation.  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (4498)..(4499)  
OTHER INFORMATION: A CAG is missing between these bases as compared  
OTHER INFORMATION: to SEQ ID NO:1.  
US-09-535-008-66

Alignment Scores:  
Pred. No.: 0.000887 Length: 5468  
Score: 150.50 Matches: 146  
Percent Similarity: 30.35% Conservative: 81  
Best Local Similarity: 19.52% Mismatches: 241  
Query Match: 3.74% Indels: 280  
DB: 4 Gaps: 31

US-09-765-272A-66 (1-763) x US-09-535-008-66 (1-5468)

Qy 177 GluAspThrGlyAspAlaTyr-----IleValProHis 187  
Db 353 GAAAGGAATGGGATGCGGTCTGAGGGGGCCATGCTGGGATGGGGCCCCCGCCAGCCCCCAT 412  
Qy 188 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAl 207  
Db 413 GGACCAGCACTCCCAAGGTTACCCC---TCGCCCCCTGGGTGGCTCTGACATGCCCTCTAG 469  
Qy 207 agluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerSerTyrAsnAl 227  
Db 470 TCCAGTT-----CCAGCAGTGGCCCGTCTCGGGGCCCCAGATGCTTC 514  
Qy 227 aAsnPro-----AlaGlnProArgLeuSerGluAsnHisAsnLe 240  
Db 515 CGGGCCAGGAGTGCCCGCTGGATGTGCTGACCCCGAGGCTTGGGGCAGCAGAACCG 574  
Qy 240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgG1 260

Db 575 GGGCCCAACCCCA---TTTAAACGAGAACCCAG-----CTGCACCA 610  
QY 260 uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPr 280  
Db 611 GCTCAGAGCTCAGATCATGCTCAAGATGCTG----- 644  
QY 280 oAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi 300  
Db 645 -----GCCAGGGGGCAGCCCTCCCGACACCTGCAGATGGC 682  
QY 300 sPheIleProTyrGlnGlnMetSerGluLeuGluLysArgIleAlaArgIle----- 317  
Db 683 GGTGCAGGGCAAGCGGCGATGCTCCGGATGCAGCAGCAGATGCCAAGCTACCTCCACC 742  
QY 317 ----- 317  
Db 743 CTCGGTGTCCGCAACAGACCGCGCTGCGCTGCGCTGCGCCCGCGCGCTCCCGG 802  
QY 318 -----IleProLeuArgTyrArgSerAsnHis----- 326  
Db 803 CCGCGCACCTCCAAATACAGAGCCCTCATGTATGGAGGGGCCACATGCCTCCCCC 862  
QY 327 -----TrpValProAspSerArgProGluGlnPro----- 336  
Db 863 AGGACCTCGGGCGTGCTCCCGCGATGCCAGGCCAGCTCTCTCGAGGGCTCCCAAGCC 922  
QY 337 -----SerProGlnSerThrProGlu----- 343  
Db 923 CTGGCTTGAAGGACCCATGTGCTGCTGCTCCCGCCACGACGCCCTCAGAGAGCTGAT 982  
QY 344 -----ProSerProSerProGlnProAlaProAsnPro----- 354  
Db 983 TCCCCCGACGACCAAGCGGCGCTTCTCCCGCGCGCTCCGCTCCACCGCGCGCTC 1042  
QY 355 -----GlnProAlaProSerAs 360  
Db 1043 GCCCGTGATCCACCGCAGACCCAGTCCCGCGGCGAGCGCGCGCGCTCCATGGT 1102  
QY 360 nProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377  
Db 1103 GCCACTGCACGAGAGCAGCGCCGATCCACCCCATCCAGAACCGCGGGGCTCGACCC 1162  
QY 378 -----TyrValPheGluAsnGlyValSerArgTyrIl 389  
Db 1163 TGTGGAGATCCTGCAGAGCGAGTACAGAGCTCGAGCTCGATCGCACCCGAATTC 1222  
QY 389 eProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys----- 405  
Db 1223 GGAACCTGAAACCTTCCCGGTCCTCGCGCGGATTTGGAACCAAGCGACCATGA 1282  
QY 405 ----- 405  
Db 1283 GCTCAAGGCCCTCAGGCTGTGTAACCTCCAGAGCGAGCTGCGCGAGGAGTGGTGTG 1342  
QY 406 -LeuAlaLysGlnGluSerLeuSerHisLysLeuAlaLys-----LysThrAspLe 423  
Db 1343 CATCGGAGGGACACAGCGCTGGAGACAGCCCTCAATGCTAAGGCTTACAGCGCAGCAA 1402  
QY 423 uProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuAlaArgIleHisGl 443  
Db 1403 CGCGCATCTCCGCGAG-----GCCCGATCACTGA 1435  
QY 443 nAspLeu-----LeuAspAsnLysGlyArgGlnValAspPheGl 456  
Db 1436 GAACCTGGAGAGCAGCAGAGATCGAGCAGAGCGCGCAAGCGCGCAGACCAAGCA 1495  
QY 456 uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val\*\* 469  
Db 1496 ATACCTCAATAGCATCTCCAGCATGCCAAGGATTTCAAGGAATATCACAGATCCGTCAC 1555  
QY 469 \*SerAspLysValLysLeuVal\*\*AspIleLeuAlaPheLeuAlaProIle----- 486

Db 1556 AGGCAAAATCCAGAGCTGACCAAGAGCGAGTGCCACGTACCATGCCAACACGAGCGGGA 1615  
QY 487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs 504  
Db 1616 GCAGAGAAAGAGAGACGAGCGATCGAGAGAGCGCATGCGGAGCTCATGCTGAAGA 1675  
QY 504 pGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAs 524  
Db 1676 TGAGGAGGGGTACCCCAAGCTCATC----- 1700  
QY 524 pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHis 544  
Db 1701 -----GACCAGAAAGAGCAAGCGCTGCGCTACCTCTTG-----CAGCAGACAGCA 1750  
QY 544 rHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyrAl 564  
Db 1751 GTACGTGCTAACCTCAGGAGCTGGTGGCGGAGCAGCAGAGGCTGCCAGGTCCCAAGGA 1810  
QY 564 aLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584  
Db 1811 GAAAAGAAAGAAAAGAAAAGAGAGCAGAGCAAGAAATGCAGAA---GGACAGACGCTGC 1867  
QY 584 aLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspAr 604  
Db 1868 CATTCGGCGGATGGC-----GAGCCTCTAGACGA 1897  
QY 604 g-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnGl 618  
Db 1898 GACCAAGCAGATGAGCGACCTCCCGGTGAGGTG-----ATCCAGTGGAGAGTGG 1948  
QY 618 ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGl 638  
Db 1949 GAAGATCTCCACAGC----- 1964  
QY 638 uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTy 658  
Db 1965 -----ACAGATGCCCCCAAGCGCGGCGAGCTGGAGGCTGCTCGAGATGAACCGGG 2017  
QY 658 rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678  
Db 2018 GTAT-----GAAGTAGCTCCGAGGTCTGAT----- 2042  
QY 678 pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGl 698  
Db 2043 -----AGTGA 2047  
QY 698 uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe 718  
Db 2048 AGAAAGTGGCTCAGAAAGAGAGAGAGAGAGAG-----GAGCAAGAGCAGCGCAGGC 2101  
QY 718 rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGl 738  
Db 2102 AGCAGAGCTCCACC-----CTGCCGCTGGAGAGAGAGAGATTCAGATCCAGACAG 2158  
QY 738 uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaGl 755  
Db 2159 CGATCAGCTCTCTGAGTGGAGCGCGCGGCACATCATTTGAGATGCCCAAGCATGTGCA 2218  
QY 755 uAspLeuLeuGlyLysIleGln 762  
Db 2219 TGATGAATATGCGTGTCCAG 2240  
RESULT 40  
US-09-535-008-1  
; Sequence 1, Application US/09535008  
; Patent No. 6465629  
; GENERAL INFORMATION:  
; APPLICANT: Wong, Alexander K.C.  
; APPLICANT: Tavtigian, Sean V.  
; APPLICANT: Teng, David H.-F.  
; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE  
; TITLE OF INVENTION: AND OTHER CANCER TYPES  
; FILE REFERENCE: 2318-259









FILE REFERENCE: 2318-259  
CURRENT APPLICATION NUMBER: US/09/535,008  
CURRENT FILING DATE: 2000-03-23  
EARLIER APPLICATION NUMBER: U.S. 60/125,806  
EARLIER FILING DATE: 1999-03-23  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 70  
LENGTH: 5480  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (75)..(5024)  
FEATURE:  
NAME/KEY: conflict  
LOCATION: (1780)  
OTHER INFORMATION: GenBank Accession No. 6455629 U29175 shows a C at this  
OTHER INFORMATION: position (position 1784 in GenBank) rather than  
OTHER INFORMATION: the G shown here.  
FEATURE:  
NAME/KEY: allele  
LOCATION: (26)  
OTHER INFORMATION: Polymorphism of either T or C in this noncoding  
OTHER INFORMATION: region.  
FEATURE:  
NAME/KEY: allele  
LOCATION: (1583)  
OTHER INFORMATION: Polymorphism of A or G resulting in a silent  
OTHER INFORMATION: mutation.  
FEATURE:  
NAME/KEY: allele  
LOCATION: (1598)  
OTHER INFORMATION: Polymorphism of T or C resulting in a silent  
OTHER INFORMATION: mutation.  
FEATURE:  
NAME/KEY: allele  
LOCATION: (1892)  
OTHER INFORMATION: Polymorphism of A or G resulting in a silent  
OTHER INFORMATION: mutation.  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (4245)..(4253)  
OTHER INFORMATION: Insertion of 9 basepairs as compared to SEQ ID  
OTHER INFORMATION: NO.1.  
US-09-535-008-70

Alignment Scores:  
Pred. No.: 0.00089 Length: 5480  
Score: 150.50 Matches: 146  
Percent Similarity: 30.35% Conservative: 81  
Best Local Similarity: 19.52% Mismatches: 241  
Query Match: 3.74% Indels: 280  
DB: 4 Gaps: 31

US-09-765-272A-66 (1-763) x US-09-535-008-70 (1-5480)

```
QY 177 GluAspThrGlyAspAlaTyr-----IleValProHis 187
Db 353 GAAAGGAATGGGGATGCGGTGAGGGGGCCATGCTGGGATGGGGCCCGCCCGCCCGCCCAT 412
QY 188 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAla 207
Db 413 GGACGAGCCTCCACAGTTACCC-----TCGCCCCCTGGGTGGCTCTGAGCATGCTCTAG 469
QY 207 aGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerTyrAsnAl 227
Db 470 TCCAGTT-----CCAGCCAGTGGCCGGTCTTCGGGGCCCGCCAGCATGCTCTTC 514
QY 227 aAsnPro-----AlaGlnProArgLeuSerGluAsnHisAsnLe 240
Db 515 CGGGCCAGAGGTGCCCGCTGGATGGTGTGCTGACCCCGCCAGGCTTTGGGGGACGAGAACCG 574
```

```
QY 240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGl 260
Db 575 GGGCCCAACCCCA---TTTAACAGAACCCAG-----CTGCACCA 610
QY 260 uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPr 280
Db 611 GCTCAGAGCTCAGATCATGGCTTACAGATGCTG-----GCTCAGAGGCGAGCCCTCCCGCAGATGGC 644
QY 280 oAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi 300
Db 645 -----GCCAGGGGCGAGCCCTCCCGCAGATGGC 682
QY 300 sPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIle----- 317
Db 683 GGTGAGGGCAAGCGCGCGATGCGCGGATGCAGCAGCAGATGCACGCTACCTCCACC 742
QY 317 ----- 317
Db 743 CTCGGTGTCCGCAACAGAGCCCGCCTGGCCCTGGCCCTGGCCCGCGCGGTCCCGG 802
QY 318 ---IleProLeuArgTyrArgSerAsnHis----- 326
Db 803 CCGGCGACCTCCAAATTACAGCAGGCTCATGTATGGAGGGCCCAACATGCTCCCCC 862
QY 327 -----TrpValProAspSerArgProGluGlnPro----- 336
Db 863 AGGACCTCGGGCGTGCCTCCCGGATGCCAGGCGAGCTCTCTGGAGGGCTCCCAAGCC 922
QY 337 -----SerProGlnSerThrProGlu----- 343
Db 923 CTGGCCTGAAGGACCATGGCGAATGCTGCTGCCCGCCAGCAGCACCCTCAGAAGCTGAT 982
QY 344 ---ProSerProSerProGlnProAlaProAsnPro----- 354
Db 983 TCCCCCGAGCAACGCGGGCGCTTCCCCCGCGCCCTCCCGCGCCCTCCCGCGCGCTC 1042
QY 355 -----GlnProAlaProSerAs 360
Db 1043 GCCCGTGTATCCACCGCAGACCCAGTCCCGCGGCGAGCCCGCCCGCCCGCCCATGGT 1102
QY 360 nProIleAspGlyLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377
Db 1103 GCCACTGCACCAAGCAGCAGCGCGCATACCCCATCCAGAACCGCGGGGCTCGACCC 1162
QY 378 -----TyrValPheGluGluAsnGlyValSerArgTyrI 389
Db 1163 TGTGGAGATCTGCAGAGCGCGAGTACAGGCTGAGGCTGCATCGCACACCAATCA 1222
QY 389 eProLalysAspLeuSerAlaGluThrAlaGlyIleAspSerLys----- 405
Db 1223 GGAACCTTGAAACCTTCCCGGGTCCCTGGCGGGGATTTCCGAAACCAAGCGACCATGA 1282
QY 405 ----- 405
Db 1283 GCTCAAGGCCCTCAGGCTGCTGAACCTTCCAGAGCGAGCTGCCCGAGAGTGTGGTGTG 1342
QY 406 -LeuAlaLysGlnGluSerHisLysLeuGlyAlaLys-----LysThrAspLe 423
Db 1343 CATGGAGGGGACACAGCGCTGGAGCAGCCCTCAATGCTAAGCCCTACAGCGCAGCAA 1402
QY 423 uProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGl 443
Db 1403 GCGCCAGTCCCTGCGCGAG-----GCCCGCATCACTGA 1435
QY 443 nAspLeu-----LeuAspAsnLysGlyArgGlnValAspPheGl 456
Db 1436 GAAGCTGGAGAGCAGCAGAGAATCGAGCAGGCGCAGCGCGCGCAGAGCAGCAGGA 1495
QY 456 uAlaLeuAspAsnLeuGluArgLeuLysAsp-----Val** 469
Db 1496 ATACCTCAATAGATTTCTCCAGATGCCAGGATTTCAAGGAATATACAGATCCGCTAC 1555
QY 469 *SerAspLysValLysLeuVal***AspIleLeuAlaPheLeuAlaProIle----- 486
```

```
1556 AGGCAAAATCCAGAGCTGACCAAGGAGTGGCCACGTACCATCCCAACAGCGAGCGGA 1615
487 QY -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs 504
1616 GCAGAGAAGAAGAACAGCGAGTGGAGAGCGGATCGGAGGCTCATGGCTGAAGA 1675
504 QY pGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAs 524
1676 TGAGGAGGGGTACCGCAAGTCTATC----- 1700
524 QY pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSe 544
1701 QY -----GACCAGAGAAGAAGAACAGCGCTGGCTCTCTCTG-----CAGCAGACAGACGA 1750
544 QY rHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAl 564
1751 GTACGTGGCTTAACCTACGAGCTGGTGGCGGACACAGAGCTGCCAGTGGCCAAAGGA 1810
564 QY aLysGluLysGlyLeuThrProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584
1811 GAAAGAGAAGAAAGAAAGAAAGAGAGGCGAAATGCAGAA---GGACAGACGCTGTC 1867
584 QY aLysGlyAlaGlnAlaLeuTyrAsnArgValLysAlaAlaLysLysValProLeuAspAr 604
1868 CATGGCGCGGATGCG-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnGl 618
604 QY -----GACCAGCAGATGACGACCTCCCGGTGAAGTG-----ATCCACGTGGAGTGG 1948
1898 GACCAGCAGATGACGACCTCCCGGTGAAGTG-----ATCCACGTGGAGTGG 1948
618 QY ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGl 638
1949 GAAGATCCTCACAGGC----- 1964
638 QY uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTy 658
1965 QY -----ACAGATGCCCCAAAGCGGCGACCTGGAGCGCTGGCTCGAGATCAACCCGGG 2017
658 QY rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678
2018 GTAT-----GAAGTAGCTCCGAGGTCTGAT----- 2042
678 QY pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGl 698
2043 QY -----AGTGA 2047
698 QY uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe 718
2048 AGAAGTGGCTCAGAAAGAGAGAGAGAGAGAG-----GAGGAAGAGAGAGAGAGAG 2101
718 QY rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGl 738
2102 AGCAGACCTCCCAACC---CTGCCCGTGGAGAGAGAGAGAGAGATTCAGATCCACAGAG 2158
738 QY uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaGl 755
2159 CGATGAGCTCTGAGGTGGAGCGCGGCGGACATCATTTAGGAATGCCAAGCAAGATGCGA 2218
755 QY uAspLeuLeuGlyLysIleGln 762
2219 TGATGAATATGGCTGTGCCAG 2240
```

## RESULT 43

US-09-535-008-68

Sequence 68, Application US/09535008

Patent No. 6465629

GENERAL INFORMATION:

APPLICANT: Wong, Alexander K.C.

APPLICANT: Tavtigian, Sean V.

APPLICANT: Teng, David H.-F.

TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE

AND OTHER CANCER TYPES

```
FILE REFERENCE: 2318-259
CURRENT APPLICATION NUMBER: US/09/535,008
CURRENT FILING DATE: 2000-03-23
EARLIER APPLICATION NUMBER: U.S. 60/125,806
EARLIER FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 68
LENGTH: 5564
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (75)..(5108)
FEATURE:
NAME/KEY: conflict
LOCATION: (1780)
OTHER INFORMATION: GenBank Accession No. 6465629 U29175 shows a C at this
OTHER INFORMATION: position (position 1784 in GenBank) rather than
OTHER INFORMATION: the G shown here.
FEATURE:
NAME/KEY: allele
LOCATION: (26)
OTHER INFORMATION: Polymorphism of either T or C in this noncoding
OTHER INFORMATION: region.
FEATURE:
NAME/KEY: allele
LOCATION: (1583)
OTHER INFORMATION: Polymorphism of A or G resulting in a silent
OTHER INFORMATION: mutation.
FEATURE:
NAME/KEY: allele
LOCATION: (1598)
OTHER INFORMATION: Polymorphism of T or C resulting in a silent
OTHER INFORMATION: mutation.
FEATURE:
NAME/KEY: allele
LOCATION: (1892)
OTHER INFORMATION: Polymorphism of A or G resulting in a silent
OTHER INFORMATION: mutation.
FEATURE:
NAME/KEY: misc feature
LOCATION: (4245)..(4340)
OTHER INFORMATION: This is a 96 base insertion compared to SEQ ID
OTHER INFORMATION: NO:1.
FEATURE:
NAME/KEY: misc feature
LOCATION: (4594)..(4595)
OTHER INFORMATION: A deletion of CAG occurs between these bases as
OTHER INFORMATION: compared to SEQ ID NO:1 (bases 4499-4501 of SEQ ID
OTHER INFORMATION: NO:1).
```

## Alignment Scores:

Pred. No.:	0.000912	Length:	5564
Score:	150.50	Matches:	146
Percent Similarity:	30.35%	Conservative:	81
Best Local Similarity:	19.52%	Mismatches:	241
Query Match:	3.74%	Indels:	280
DB:	4	Gaps:	31

US-09-765-272A-66 (1-763) x US-09-535-008-68 (1-5564)

```
QY 177 GluAspThrGlyAspAlaTyr-----IleValProHis 187
Db 353 GAAAGGAATGGGATGCGGTTCAGGGGGCCATGCTGGATGGGCGCCCGCCAGCCCAT 412
QY 188 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAl 207
Db 413 GGACCAGCACTCCCAAGGTTACCCC---TCGCCCTGGGTGGTCTGAGCATGCTCTAG 469
QY 207 aGluAlaTyrTrpAsnGlyAsnGlnGlySerArgProSerSerSerSerTyrAsnAl 227
```

Db 470 TCCAGTT-----CCAGCCAGTGGCCGCTCTCGGGGCCCCAGATGCTCTTC 514  
 QY 227 aAsnPro-----AlaGlnProArgLeuSerGluAAsnHisLe 240  
 Db 515 CGGCCAGAGGTCCCGCTGGATGGTGTGACCCCGAGGCTTGGGGCAGCAACCG 574  
 QY 240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnHisLeuArgG 260  
 Db 575 GGGCCCAACCCCA---TTTAAACCAAGAACAG-----CTCCACCA 610  
 QY 260 uLeuTyrAlaLysProLeuSerGluAArgHisValGluSerAspGlyLeuLeuPheAsp 280  
 Db 611 GCTCAGAGCTCAGATCAGGCTACAGATGCTG----- 644  
 QY 280 oAlaGlnInleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyr 300  
 Db 645 -----GCCAGGGGCGAGCCCTCCCGACCACTCCAGATGGC 682  
 QY 300 sPheIleProTyrGluGlnMetSerGluLeuGlyLysArgIleAlaArgIle----- 317  
 Db 683 GGTGCAGGCAAGCGCGATGCCCGGATGCAGCAGATGCCAACGCTACCTCCACC 742  
 QY 317 ----- 317  
 Db 743 CTGGTGTCCCAACAGACCCGGCCCTGGCCCTGGCCCTCCCGCCCGGGTCCCGG 802  
 QY 318 -----IleProLeuArgTyrArgSerAsnHis----- 326  
 Db 803 CCGGGCACCTCAAATACAGCAGCCTCATGGTATGGAGGGCCCAACATGCTCCCTCC 862  
 QY 327 -----TrpValProAspSerArgProGluGlnPro----- 336  
 Db 863 AGGACCCTCGGGCTGCGCCCGGATGCCAGCCAGCTCTCTGAGGGCCCTCCCAAGCC 922  
 QY 337 -----SerProGlnSerThrProGlu----- 343  
 Db 923 CTGGCTTGAGGACCCATGGCGAATGTGCTGCCCCCAGCAGCACCCCTCAGAACTGAT 982  
 QY 344 -----ProSerProSerProGlnProAlaProAsnPro----- 354  
 Db 983 TCCCGCCGAGCAAGCGGCGCCCTTCCCGCGCGCCCTGCGGCCCGCCGCTC 1042  
 QY 355 -----GlnProAlaProSerAs----- 360  
 Db 1043 GCCCGTGATGCCACCGCAGACCCAGTCCCGCGGCGAGCGCGCCCGCCCGCTG 1102  
 QY 360 nProIleAspGlyLeuValLysGluAlaValArgLysValGlyAspGly----- 377  
 Db 1103 GCCACTGACCAAGACAGCGCCGATCACCCCATCCAGAACCGCGGGGCTCGACCC 1162  
 QY 378 -----TyrValPheGluGluAsnGlyValSerArgTyrI 389  
 Db 1163 TGTGGAGATCTCGCAGGAGCGGAGTACAGGCTGCAGCTCGCATCGCACCCGAATCA 1222  
 QY 389 eProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys----- 405  
 Db 1223 GGAACCTTGAACACCTTCCCGGGTCTCGCGGGGATTTGCAACCAAGCAGCCATCA 1282  
 QY 405 ----- 405  
 Db 1283 GCTCAAGGCCCTCAGGTGTGAATCTCCAGAGGCGAGTGGCCAGGAGGTGGTGGTGG 1342  
 QY 406 -LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys-----LysThrAspLe 423  
 Db 1343 CATCGGGAGGACACAGCGCTGGAGACAGCCCTCAATGTATAGGCTACACAGCGCAGCA 1402  
 QY 423 uProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHis 443  
 Db 1403 GCGCCAGTCCCTGCGCGAG-----GCCCGCATCACTGA 1435  
 QY 443 nAspLeu-----LeuAspAsnLysGlyArgGlnValAspPheG 456  
 Db 1436 GAAGCTGGAGAGCAGCAGAGATCGACGAGGCGCAAGCGCCGCGCAGCAGCA 1495

QY 456 uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val\*\* 469  
 Db 1496 ATACCTCAATAGCATCTCCAGCATCCCAAGGATTTCAAGGAATATACAGATCCGTCAC 1555  
 QY 469 \*SerAspLysValLysLeuVal\*\*AspIleLeuAlaPheLeuAlaProIle----- 486  
 Db 1556 AGGCAAAATCCAGAACTGACCAAGCAGTGGCGCAGTACCATGCTCCCAACGAGCGGGA 1615  
 QY 487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs 504  
 Db 1616 GCAGAGAAAGAGAACGCGGATCGAGAAGGAGCGCATCGGAGGCTCATGGCTGAAGA 1675  
 QY 504 pGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAs 524  
 Db 1676 TGAGGAGGGGTACCGCAAGCTATC----- 1700  
 QY 524 pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSe 544  
 Db 1701 -----GACCAAGAAAGACAAAGCGCTGGCTCTCTTG---CAGCAGCAGCAGGA 1750  
 QY 544 rHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAl 564  
 Db 1751 GTACGTGGCTAACCTCAGGAGCTGTGGCGGCAGCACAGGCTGCCAGGTCCCAAGGA 1810  
 QY 564 aLysGluLysGlyLeuThrProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584  
 Db 1811 GAAAAAGAGAAAAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1867  
 QY 584 aLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspAr 604  
 Db 1868 CATTGGCGCGGATGGC-----GAGCCTCTAGACGA 1897  
 QY 604 g-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnGl 618  
 Db 1898 GACCACCCAGATGAGCGACTCCCGGTGAGGTG-----ATCCAGCTGGAGATGG 1948  
 QY 618 ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGl 638  
 Db 1949 GAAGATCTCTCACAGG----- 1964  
 QY 638 uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTy 658  
 Db 1965 -----ACAGATGCCCCCAAGCCCGGCGAGGCTGGCTCGAGATGAACCCGGG 2017  
 QY 658 rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678  
 Db 2018 GTAT-----GAAGTAGTCCGAGTCTGAT----- 2042  
 QY 678 pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGl 698  
 Db 2043 -----AGTGA 2047  
 QY 698 uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe 718  
 Db 2048 AGAAAGTGGCTCAGAAAG 2101  
 QY 718 rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGl 738  
 Db 2102 AGCACAGCTCTCCACC---CTGCCCTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2158  
 QY 738 uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaGl 755  
 Db 2159 CGATGAGCTCTGAGGTGAGCGCGGCGACATCATTTGAGATGTCACCAAGCAGATGTCGA 2218  
 QY 755 uAspLeuLeuGlyLysIleGln 762  
 Db 2219 TGATGAATATGCGTGTCCAG 2240

## RESULT 44

US-09-535-008-64  
 ; Sequence 64, Application US/09535008  
 ; Patent No. 6465629



QY	456	uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp	-----Val** 469
Db	1496	ATACCTCAATAGCATTTCTCAGCATGCCAAGGATTTCAAGAAATATACAGATCCGTAC	1555
QY	469	*SerAspLysValLeuVal***AspIleLeuAlaPheLeuAlaProIle	-----486
Db	1556	AGGCAAAATCCAGAACTGACCAAGGCAGTGGCCGCTACCATGCCAAACGAGCGGGA	1615
QY	487	-----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs	504
Db	1616	GCAGAAGAAGAACGACGGATCGAAGAGGAGCGCATCGGAGGCTCATGCTGAAGA	1675
QY	504	pGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAs	524
Db	1676	TGAGGAGGGGTACCGCAAGCTCATC	-----1700
QY	524	pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHis	544
Db	1701	-----GACCAGAAGAAGCAAGCGCTCGGCTACCTCTTG	-----CAGCAGACGACACGA 1750
QY	544	rHisTriPleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAl	564
Db	1751	GTACGTGGCTAACTCAGCGAGTGGTGGCGGACGACACAGGCTGCCAGGTCCCAAGGA	1810
QY	564	AlysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAl	584
Db	1811	GAATAAGAGAAAAGAAAAGAAAGAGGAGGAGAAATGCGAA	-----GGACAGACGCGCTGC 1867
QY	584	AlysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspAr	604
Db	1868	CATTGGCCGGATGGC	-----GAGCCTCTAGACGA 1897
QY	604	g-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnGln	618
Db	1898	GACCAGCCAGATGAGCGACCTCCCGGTGAAGNG	-----ATCCAGCTGGAGAGTGG 1948
QY	618	ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGln	638
Db	1949	GAAGATCCTCAAGGC	-----1964
QY	638	uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyr	658
Db	1965	-----ACAGATGCCCCCAAAACCGCGGACGTGGAGGCGCTGGCTCAGATGAACCCGGG	2017
QY	658	rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs	678
Db	2018	GTAT-----GAAGTAGCTCCGAGTCTGTAT	-----2042
QY	678	pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGln	698
Db	2043	-----AGTGA	2047
QY	698	uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe	718
Db	2048	AGAAATGGCTCAGAAGAGAGAGAGAGAGGAG	-----GAGGAAGAGACGCGCAGGC 2101
QY	718	rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGln	738
Db	2102	AGCACAGCCTCCACC-----CTGCCCTGGAGGAGAGAAGAATTCAGATCCACAG	2158
QY	738	uGluProGlnValGluThrGluLys	-----ValGluGluLysLeuArgGluAlaGln 755
Db	2159	CGATGACGCTCTGAGGTGACCGCGGCACATCATTTAGAAATGCCAAGCAATCTCGA	2218
QY	755	uAspLeuLeuGlyLysIleGln	762
Db	2219	TGATGAATATGGCGGTGCCAG	2240

```

1  GENERAL INFORMATION:
2  APPLICANT: Wong, Alexander K.C.
3  APPLICANT: Tavtigian, Sean V.
4  APPLICANT: Teng, David H.-F.
5  TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE
6  TITLE OF INVENTION: AND OTHER CANCER TYPES
7  FILE REFERENCE: 2318-259
8  CURRENT APPLICATION NUMBER: US/09/535,008
9  CURRENT FILING DATE: 2000-03-23
10 EARLIER APPLICATION NUMBER: U.S. 60/125,806
11 EARLIER FILING DATE: 1999-03-23
12 NUMBER OF SEQ ID NOS: 77
13 SOFTWARE: PatentIn Ver. 2.0
14 SEQ ID NO 76
15 LENGTH: 5573
16 TYPE: DNA
17 ORGANISM: Homo sapiens
18 FEATURE:
19 NAME/KEY: CDS
20 LOCATION: (75)..(5117)
21 FEATURE:
22 NAME/KEY: conflict
23 LOCATION: (1780)
24 OTHER INFORMATION: GenBank Accession No. 645629 U29175 shows a C at this
25 OTHER INFORMATION: position (position 1784 in GenBank) rather than
26 OTHER INFORMATION: the G shown here.
27 FEATURE:
28 NAME/KEY: allele
29 LOCATION: (26)
30 OTHER INFORMATION: Polymorphism of either T or C in this noncoding
31 OTHER INFORMATION: region.
32 FEATURE:
33 NAME/KEY: allele
34 LOCATION: (1583)
35 OTHER INFORMATION: Polymorphism of A or G resulting in a silent
36 OTHER INFORMATION: mutation.
37 FEATURE:
38 NAME/KEY: allele
39 LOCATION: (1598)
40 OTHER INFORMATION: Polymorphism of T or C resulting in a silent
41 OTHER INFORMATION: mutation.
42 FEATURE:
43 NAME/KEY: allele
44 LOCATION: (1892)
45 OTHER INFORMATION: Polymorphism of A or G resulting in a silent
46 OTHER INFORMATION: mutation.
47 FEATURE:
48 NAME/KEY: misc feature
49 LOCATION: (4245)..(4349)
50 OTHER INFORMATION: Addition of 105 basepairs compared to SEQ ID NO:1.
51 FEATURE:
52 NAME/KEY: misc feature
53 LOCATION: (4603)..(4604)
54 OTHER INFORMATION: Deletion of CAG between these two basepairs as
55 OTHER INFORMATION: compared to SEQ ID NO:1 (deletion of basepairs
56 OTHER INFORMATION: 4499-4501 of SEQ ID NO:1).
57 US-09-535-008-76

```

Alignment Scores:	
Pred. No.:	0.000914
Score:	150.50
Percent Similarity:	30.35%
Best Local Similarity:	19.52%
Query Match:	3.74%
DB:	4
Length:	5573
Matches:	146
Conservative:	81
Mismatches:	241
Indels:	280
Gaps:	31

US-09-765-272A-66 (1-763) x US-09-535-008-76 (1-5573)

Qy	177	GluAspThrGlyAspAlaTyr-----IleValProHis	187
Db	353	GAAAGGAATGGGGATGGCTCAGGGGGGCATCTGGGATGGGGCCGCCAGCCCAT	412
Qv	188	Gly-AspHisTyrHisTvrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaAl	207

413 GACACGACACTCCCAAGGTTACCCC---TGCCCCCTGGGTGGCTCTGAGCATGCTCTAG 469  
207 agluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerSerTyrAsnAl 227  
470 TCCAGTT-----CCAGCAGTGGCCGCTCTTCGGGGCCCGCAGATGCTCTTC 514  
227 aAsnPro-----AlaGlnProArgLeuSerGluAsnHisAsnLe 240  
515 CGGGCCAGGAGTGGCCCGCTGGATGGTGTGACCCCGCCAGGCTTGGGGCAGCAACCG 574  
240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgG 260  
575 GGGCCCCAACCCCA---TTTAAACGAGAACCCAG-----CTGCACCA 610  
260 uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPr 280  
611 GCTCAGAGCTCAGATCATGCTTACAGAGTCTG----- 644  
280 oAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi 300  
645 -----CCAGGGGGCAGCCCTCCCGACCACTGCAGATGGC 682  
300 sPheIleProTyrGluGlnMetSerGluLeuLysArgIleAlaArgIle----- 317  
683 GGTGACGGGCAAGCGCGATGCCGGGATGCAGCAGATGCCAACGCTACCTCCAC 742  
317 ----- 317  
743 CTCGGTGTCCGCAACAGACCCGCGCTGCTGGCCCTGGCCCGCCCGGGTCCCGG 802  
318 -----IleProLeuArgTyrArgSerAsnHis----- 326  
803 CCGGCACCTCCAAATACAGCAGGCTCATGTATGGAGGGCCCAACATGCCTCCCCC 862  
327 -----TrpValProAspSerArgProGluGlnPro----- 336  
863 AGGACCTCGGGGTGCCCCCGGGATGCCAGGCCAGCTCTTGAGGGCTCCCAAGCC 922  
337 -----SerProGlnSerThrProGlu----- 343  
923 CTGGCTGAAGCACCCTATGCGAATGCTGTGCCCCCAGCAGCACCCTCAGAACTCAT 982  
344 -----ProSerProSerProGlnProAlaProAsnPro----- 354  
983 TCCCCGCGACCAACGGCGCCCTTCCCCCGCGCCCTCTGCGCTCCACCGCGCGCTC 1042  
355 -----GlnProAlaProSerAs 360  
1043 GCGCGTATGCCACGACAGCCAGTCCCGGGCAGCGCCCGCCCGCCATGGT 1102  
360 nProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377  
1103 GGCACCTGCACAGAGCAGCGCGATCACCCCATCCAGAAAGCGCGGGCTCGACCC 1162  
378 -----TyrValPheGluGluAsnGlyValSerArgTyrI 389  
1163 TGTGGAGATCTTCAGAGAGCGCAGTACAGGCTGCGAGCTGCGATCGCACACCGAATCA 1222  
389 eProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys----- 405  
1223 GGNACTTGAAACCTTCCCGGGTCCCTGGCGGGGATTTGCGAACCAAGCGACCAATGA 1282  
405 ----- 405  
1283 GCTCAAGGCGCTCAGGCTGCTGAACCTCCAGAGCAGCTGCGCCAGAGGTGGTGTG 1342  
406 -LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys-----LysThrAspLe 423  
1343 CATCGGAGGGACACAGCGCTGAGACAGCCCTCAATGTCTAAGCCCTTACAGCGCACCA 1402  
423 uProSerAspArgGluPheTyrAsnLysAlaTyrAspLeuAlaArgIleHisG 443

1403 GCGCCAGTCCCTGCGCGAG-----GCCGCGATCACTGA 1435  
443 nAspLeu-----LeuAspAsnLysGlyArgGlnValAspPheG 456  
1436 GAAGCTGGAGAGCAGCAGAGATCGAGCAGAGCGCAAGCGCCGAGAGCAGCAGGA 1495  
456 uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val\*\* 469  
1496 ATACCTCAATGACTTCTCCAGCATGCCAAGGATTTCAAGGAATATCAAGATCCGCTCAC 1555  
469 \*SerAspLysValLysLeuVal\*\*AspIleLeuAlaPheLeuAlaProIle----- 486  
1556 AGCAAAATCCAGAGCTGACCAAGCAGTGCCTACCTACCTCCCAACACGAGCGGGA 1615  
487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs 504  
1616 GCAGAAAGAGAGAACGAGCGGATCGAAGAGCGCATGCGAGGCTCATGGCTGAAGA 1675  
504 pGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAs 524  
1676 TGAGGAGGGTACCGCAAGCTCATC----- 1700  
524 pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSe 544  
1701 -----GACCAGAAAGAGCAAGCGCTGCTACCTCTTG---CAGCAGACGACGGA 1750  
544 rHisTriPheLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAl 564  
1751 GTACGTGGCTAACCTCAGGAGCTGGTGGGCGAGCACAAGGCTGCCAGGTGCCAAGGA 1810  
564 aLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584  
1811 GAAAAGAGAGAAAGAAAGAGAGCAGCAAAATGCAGAA---GGACAGACGCTGC 1867  
584 aLysGlyAlaGluAlaIleTyrAsnArgValLysAlaLysLysValProLeuAspAr 604  
1868 CATGGGCGGATGTC-----GAGCCTCTAGACGA 1897  
604 g-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnG 618  
1898 GACCAGCCAGATGACGACCTCCCGGTGAAGTG-----ATCCAGTGGAGAGTGG 1948  
618 ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspG 638  
1949 GAAGATCTCACAGC----- 1964  
638 uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTy 658  
1965 -----ACAGATGCCCAAGCGCGCAGCTGGAGGCTGCTCGAGATGAACCCGGG 2017  
658 rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678  
2018 GTAT-----GAAGTAGTCTCCGAGGTCTCAT----- 2042  
678 pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerG 698  
2043 -----AGTGA 2047  
698 uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe 718  
2048 AGAAGTGGCTCAGAGAGAGAGAGAGAGAGAG-----GAGGAAGAGCAGCGCCGAGC 2101  
718 xGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerG 738  
2102 AGCAGCGCTCCACC-----CTGCGCGTGGAGAGAGAGAGATTCAGATCCAGACAG 2158  
738 uGluProGlnValThrGluLys-----ValGluGluLysLeuArgGluAlaG 755  
2159 CGATGACGCTCTCTGAGTGGAGCGCGCGGCACATCATTTGAGAAATGCCAAGATGTGCA 2218  
755 uAspLeuLeuGlyLysIleGln 762  
2219 TGATCAATATGCGGTGCTCCAG 2240

RESULT 46  
 US-09-535-008-72  
 ; Sequence 72, Application US/09535008  
 ; Patent No. 6465629  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wong, Alexander K.C.  
 ; APPLICANT: Tavtigian, Sean V.  
 ; APPLICANT: Teng, David H.-F.  
 ; TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE  
 ; TITLE OF INVENTION: AND OTHER CANCER TYPES  
 ; FILE REFERENCE: 2318-259  
 ; CURRENT APPLICATION NUMBER: US/09/535,008  
 ; CURRENT FILING DATE: 2000-03-23  
 ; EARLIER APPLICATION NUMBER: U.S. 60/125,806  
 ; EARLIER FILING DATE: 1999-03-23  
 ; NUMBER OF SEQ ID NOS: 77  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 72  
 ; LENGTH: 5576  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (75)..(5120)  
 ; FEATURE:  
 ; NAME/KEY: conflict  
 ; LOCATION: (1780)  
 ; OTHER INFORMATION: GenBank Accession No. 6465629 U29175 shows a C at this  
 ; position (position 1784 in GenBank) rather than  
 ; the G shown here.  
 ; FEATURE:  
 ; NAME/KEY: allele  
 ; LOCATION: (26)  
 ; OTHER INFORMATION: Polymorphism of either T or C in this noncoding  
 ; region.  
 ; FEATURE:  
 ; NAME/KEY: allele  
 ; LOCATION: (1583)  
 ; OTHER INFORMATION: Polymorphism of A or G resulting in a silent  
 ; mutation.  
 ; FEATURE:  
 ; NAME/KEY: allele  
 ; LOCATION: (1598)  
 ; OTHER INFORMATION: Polymorphism of T or C resulting in a silent  
 ; mutation.  
 ; FEATURE:  
 ; NAME/KEY: allele  
 ; LOCATION: (1892)  
 ; OTHER INFORMATION: Polymorphism of A or G resulting in a silent  
 ; mutation.  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (4245)..(4349)  
 ; OTHER INFORMATION: This is a 105 basepair insertion as compared to  
 ; the other information: SEQ ID NO:1.  
 ; OTHER INFORMATION: SEQ ID NO:1.

Alignment Scores:  
 Pred. No.: 0.000915 Length: 5576  
 Score: 150.50 Matches: 146  
 Percent Similarity: 30.35% Conservative: 81  
 Best Local Similarity: 19.52% Mismatches: 241  
 Query Match: 3.74% Indels: 280  
 DB: 4 Gaps: 31

US-09-765-272A-66 (1-763) x US-09-535-008-72 (1-5576)

Qy 177 GluAspThrGlyAspAlaTyr-----IleValProHis 187  
 Db 353 GAAGAGTGGGATGGCGTTCAGGGGGCGCATGCTGGGATGGGGCCCGCCCGCCGAT 412  
 Qy 188 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAla 207

Db 413 GGACGAGCAGCTCCCAAGAGTTACCCC---TCGCCCCCTGGGTGGCTCGAGCATCCTCTAG 469  
 Qy 207 aGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerSerTyrAsnAl 227  
 Db 470 TCCAGTT-----CCAGCCAGTGGCGCTCTCGGGGCCCCAGATGCTTTC 514  
 Qy 227 aasnPro-----AlaGlnProArgLeuSerGluAsnHisAsnLe 240  
 Db 515 CGGGCCAGGAGTGGCCCGCTGGATGGTGTGACCCCGAGGCTTGGGGCAGCAGAACCG 574  
 Qy 240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnLysSerSerLeuLeuArgL 260  
 Db 575 GGGCCCAACCCCA---TTTAACCAAGAACAG-----CTGCACCA 610  
 Qy 260 uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuLeuPheAspPr 280  
 Db 611 GCTCAGAGCTCAGATCATGGCTACAGATGCTG----- 644  
 Qy 280 oAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi 300  
 Db 645 -----GCCAGGGGGCAGCCCTCCCGCAGCATGCCACCTGCAGATGGC 682  
 Qy 300 sPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIle----- 317  
 Db 683 GGTGAGGCAAGCGCGCATGCCGGGATGCAGCAGCAGATGCCAACGCTACTCTCCACC 742  
 Qy 317 ----- 317  
 Db 743 CTCGGTGTCCGCAACAGGACCGCGCTGGCCCTGGCCCTGGCCCGCGGGTCCCGG 802  
 Qy 318 ----IleProLeuArgTyrArgSerAsnHis----- 326  
 Db 803 CCGGCACCTTCCAAATTACAGCAGGCTCATGTATGGAGGGCCCAACATGCTCTCCCC 862  
 Qy 327 -----TrpValProAspSerArgProGluGlnPro----- 336  
 Db 863 AGGACCTCGGGCGTGCCTCCCGGATGCCAGCCAGCTCTCTGGAGGGGCTCCCAAGCC 922  
 Qy 337 -----SerProGlnSerThrProGlu----- 343  
 Db 923 CTGGCTGAAGACCCATGCGGATCTGCTGCTGCCCCACAGCAGCAGCCCTCAGAGTGTAT 982  
 Qy 344 ----ProSerProSerProGlnProAlaProAsnPro----- 354  
 Db 983 TCCCGCGCAGCAACCGGCGCTTCCCGCGCCCTCCCGCTCCACCCCGCGCGCTC 1042  
 Qy 355 -----GlnProAlaProSerAs 360  
 Db 1043 GCCCGTGTATGCCACCGCAGACCCAGTCCCGCGGCGAGCCCGCCCGCCCGCTGGT 1102  
 Qy 360 nProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377  
 Db 1103 GCCATGTACCAGACAGCAGCGCATCCCCCATCCAGAGCGCGGGCGCTCGACCC 1162  
 Qy 378 -----TyrValPheGluGluAsnGlyValSerArgTyrI 389  
 Db 1163 TGTGGAGATCTTCAGGAGCGGAGTACAGGCTGCGAGCTCGCATCGCACACCGAATTCA 1222  
 Qy 389 eProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys----- 405  
 Db 1223 GGAACCTTGAACCTTCCCGGGTCCCTGGCGGGGATTTCCGAACCAAGCAGCATTGA 1282  
 Qy 405 ----- 405  
 Db 1283 GCTCAAGGCCCTCAGGCTGCTGAACCTTCAGAGCGAGCTCGCCGAGGAGGTGGTGTG 1342  
 Qy 406 -LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys-----LysThrAspLe 423  
 Db 1343 CATCGCGGGACACAGCGCTGGAGACAGCCCTCAATGCTAAGCCCTACAGCGCAGCAA 1402  
 Qy 423 uProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGL 443



1403	GGCGCAGTCCCTCGCGGAG----	DB	-----GCCGCGATCACTGA	1433
443	naSpLeu-----	2Y	-----LeuAspAenLysGlyArgGlnValAspPheG1	456
1436	GAAGCTGGAGAGCAGCAGAGATCGAGCAGGCGCAAGCGCGCGCAGAACCCAGGA	DB		1495
456	uAlaLeuAspAenLeuLeuGluArgLeuLysAsp-----	2Y	-----Val**	469
1436	ATACCTCATAGCATTTCTCCAGCATGCCAAGGATTTCAAGGAATATCACAGATCCCGTCA	DB		1555
469	*SerAspLysValLysLeuVal**AspIleLeuAlaPheLeuAlaProIle-----	2Y	-----	486
1556	AGGCAAAATCCAGAGAGCTGACCAAGGCAGTGGCCCGTACCTACCTCCCAACACAGGAGCGGGA	DB		1615
487	-----ArgHisProGluArgLeuGlyLysProAenAlaGlnIleThrTyrThrAspAs	2Y		504
1616	GCAGAAGAAAGAGACGAGCGATCGAGAAGGAGCGCATCGGAGGCTCATGCTGAAGA	DB		1675
504	pGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAs	2Y		524
1676	TCAGAGGGGTACCGCAAGCTCATC-----	DB		1700
524	pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSe	2Y		544
1701	-----GACCAAGAAAGAGCAAGCGCTGGCCTACCTCTTG-----	DB		1750
544	rHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyrAl	2Y		564
1751	GTACGTGGCTAACCTCACGGAGCTGGTGGCGGACGACACAGGCTGCCAGTCCCAAGGA	DB		1810
564	aLysGluLysGlyLeuThrProSerThrAspHisGlnAspSerGlyAsnThrGluAl	2Y		584
1811	GAAAAAGAGAAAAAGAAAAAGAAAGAGGACAGAAAATGCAGAA-----	DB		1867
584	aLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspAr	2Y		604
1868	CATTGGCGCGATGGC-----	DB		1897
604	G-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnG1	2Y		618
1898	GACCAAGCCAGATGAGCGACCTCCCGTGAAGTG-----	DB		1948
618	ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspG1	2Y		638
1949	GAAGATCCCTCACAGGC-----	DB		1964
638	uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTy	2Y		658
1965	-----ACAGATGCCCCCAAGCCCGCAGCTGGAGGCGCTCGAGATGAACCCGGG	DB		2017
658	rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs	2Y		678
2018	GTAT-----GAAGTAGCTCCGAGTCTGAT-----	DB		2042
678	pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerG1	2Y		698
2043	-----AGTGA	DB		2047
698	uGluLysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSe	2Y		718
2048	AGAAAGTGGCTCAGAAAGAGAGAGAGAGAGAGAG-----	DB		2101
718	rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerG1	2Y		738
2102	AGCACAGCTTCCACC-----CTGCCCTGGAGGAGAGAGAGAGATTCAGATCCAGACAG	DB		2158
738	uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaG1	2Y		755
2159	CGATGACGCTCTGAGGTGACGCGCGGCACATCATTTGAGAAATGCCAAGCAAGATGTGGA	DB		2218
755	uAepLeuLeuGlyLysIleGln	762		
2219	TGATGAATATGGCGTGTCCACG	2240		

```

RESULT 47
US-08-446-137B-1
Sequence 1, Application US/08446137B
Patent No. 6162903
GENERAL INFORMATION:
APPLICANT: Trowern, Angus R.
APPLICANT: Atkinson, Anthony P.
APPLICANT: Murphy, Jonathan P.
APPLICANT: Laurence, Oliver S.
APPLICANT: Dugleby, Clive J.
TITLE OF INVENTION: IMMUNOGLOBULIN BINDING PROTEINS DERIVED
TITLE OF INVENTION: FROM L PROTEIN AND THEIR USES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,137B
FILING DATE: 22-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 100084.406
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3279 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Peptostreptococcus asaccharolyticus
STRAIN: 1018
FEATURE:
NAME/KEY: exon
LOCATION: 103..3186
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /codon_start= 280
OTHER INFORMATION: /product= "mature protein L"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: CDS
LOCATION: 103..3186
OTHER INFORMATION: /codon_start= 103
OTHER INFORMATION: /product= "immature protein L"
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 280..3183
OTHER INFORMATION: /codon_start= 280
OTHER INFORMATION: /product= "mature protein L"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 208..279
US-08-446-137B-1

```

## Alignment Scores:

Pred. No.: 0.000437 Length: 3279  
 Score: 150.00 Matches: 179  
 Percent Similarity: 31.86% Conservative: 109  
 Best Local Similarity: 19.80% Mismatches: 292  
 Query Match: 3.73% Indels: 325  
 DB: 3 Gaps: 47

US-09-765-272A-66 (1-763) x US-08-446-137B-1 (1-3279)

QY 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21  
 DB 938 AGTTATCTGAAGAAACACAG---AACGAAAGAAAGATTCAATCAAGCTTACT 994  
 QY 22 TyrIleAspGlyAspGlnAlaGlyGlnLysAlaGlu----- 33  
 DB 995 TA-ATCTTTCAGATGGAGACACACAAAATGCAGAAATTCAAAGAACATTCGCAAAAGCA 1053  
 QY 34 -----AsnLeuThrProAspGluValSerLysArgGluGlyIleAsnAlaGlu 49  
 DB 1054 GTATCAGATGCTTACGCTTACGAGATGCTTTAAAGAAACACACACGGA----- 1101  
 QY 50 Gln\*\*\*ValIleLysIleThrAspGlnGlyTyrValThrSerHisGlyAspHisTyrHis 69  
 DB 1102 GAATATACGTGAGACGTTGCAGATAAAGCTTAACTTTAAAT-----ATT 1146  
 QY 70 TyrTyrAsnGlyLys-----ValProTyrAspAlaIle---IleSerGlu 83  
 DB 1147 AAATTCGCTGTAAAGAAAGAAACACAGAGAACCAAGAGAGATTACATCAAGTT 1206  
 QY 84 GluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleValAsnGluIle 103  
 DB 1207 AACTTAATCTTGCAGATGGAAAGACACAA-----ACAGCAGAAATTC 1248  
 QY 104 LysGlyGlyTyr-----ValIleLys 110  
 DB 1249 AARGAACATTGTGAAGAGCAACACAGAAAGCTTTATGCTTATGCAGACTTTATAGCAAAA 1308  
 QY 111 ValAsnGlyLysTyrTyrValTyrLeuLysAspAlaAlaHisAlaAspAsnIleArg--- 129  
 DB 1309 GAAATGCGAATATACAGCAGACTTAGAGATGGTGGAAACACATCAATTAATTT 1368  
 QY 130 -----ThrLysGluGluIle----- 134  
 DB 1369 GCTGMAAAGAAACACCCAGAACACAGAGAACCAAGAAAGATTACAAATCAAAGTT 1428  
 QY 135 -----LysArgGlnLysGlnGluArgSerHisAsnHisAsnSer 147  
 DB 1429 AACTTAATCTTTCAGATGGAAAGATACAAACAGCAGAAATTCAAAGGAACATTTGAAGAA 1488  
 QY 148 ArgAlaAspAsnAlaValAla-----AlaAlaArgAlaGlnGlyArgTyrThr 163  
 DB 1489 GCAACAGCAAAAGCTTATGCTTATGCAAACTTATAGCAAAAGAAATGCGCAATATACA 1548  
 QY 164 ThrAspAspGlyTyrIlePheAsnAlaSerAspIlelleGluAspThrGlyAspAlaTyr 183  
 DB 1549 GCAGAC-----TTAGAAGATGGTGGAAACACACAAATC 1578  
 QY 184 IleVal-----ProHisGlyAspHisTyrHisTyrIleProLys 196  
 DB 1579 AACATTAATTTGCTGAAAGAAACACCAAGAAACACCAAGAA-----CCAAA 1629  
 QY 197 AsnGluLeuSerAlaSer----- 202  
 DB 1630 GAAGAGATTACATCAAGTTAACTTAATCTTTGCGATGGAAAGAAACACCAAGCAGAA 1689  
 QY 203 -----GluLeuAlaAlaGluAlaTyr----- 210  
 DB 1690 TTCAAAGGNACATTGGAAGAACCAACAGAGAGCTTACAGATATGCAGACTTATTAGCA 1749  
 QY 211 ---TrpAsnGlyLysGlnGlySerArgProSerSerSerSerTyrAsnAlaAsn--- 228  
 DB 1750 AAAGTAATTTGGTGAATAC---ACAGCAGACTTTAGAAGATGGCGGATACACTATCAACATC 1806

QY 229 -----ProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrPro 244  
 DB 1807 AAATTTGCTGAAAAAGAACCAACCA-----GGCGAAAAATCCAGGAATCCACAATTGATGAA 1860  
 QY 245 ThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLys 264  
 DB 1861 TGGTTATTAAAGAAAT-----GCTAAAGAAAGAACCAATCAAA 1896  
 QY 265 ProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIleThr 284  
 DB 1897 GAATTAAGAACAGCAATCACTTCTGAT---TTATACTTCAGCTTAATCAAT---AAA 1950  
 QY 285 SerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyr 304  
 DB 1951 GCAAAAACAGTTGAAGCGTA----- 1971  
 QY 305 GluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrArgSer 324  
 DB 1972 -----GAAGCATTTAAAGAACGAAATCTTAAAGCA----- 2001  
 QY 325 AsnHisTyrValProAspSerArgProGluGlnProSerProGlnSerThrProGluPro 344  
 DB 2002 -----CACGCTGGAGAAACACACCAATTAAGATGGATATGCACATATGAAGAA 2055  
 QY 345 SerProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGlu 364  
 DB 2056 GCA-----GAA 2061  
 QY 365 LysLeuValLysGluAlaValArgLys-----ValGlyAspGlyTyrValPheGluGlu 382  
 DB 2062 GCAGCAGCTAAAGAGCTTTGAAAAATGATGTTTAAACACGCATACGAAATAGTTCAA 2121  
 QY 383 AsnGlyValSerArgTyrIle-----Pro 390  
 DB 2122 GTGTCAGACGGAAGATACTACTATGATTAAGATTGAGTGCAGACGAAAGAAACCA 2181  
 QY 391 AlaLysAspLeuSerAlaGluThrAlaAlaGlyIle-----AspSerLysLeu 406  
 DB 2182 GGTGAAGAC---ACTCCAGAGTTCAAGAGAGTTTACGCACTTACGAAGAGCAGACGA 2238  
 QY 407 AlaLysGlnGluSerLeuSerHis---LysLeuGlyAlaLysLysThrAspLeuProSer 425  
 DB 2239 GCAGCTAAAGAGCAATTAAGAAAGATATAAGTAACTTAACTATGCATACGAAGTATCAAGGT 2298  
 QY 426 SerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeu 445  
 DB 2299 GCAGCGGAAGATACTAC-----TATGTATTAAAAATCGAAGATAAAGAAAGATGAA 2349  
 QY 446 Leu-----AspAsnLysGlyArgGlnValAsp-----Phe 455  
 DB 2350 CAACACAGGTGAAGAACCCAGCGAAACCCAGGAATCAATTTGATGATGTTATTAAAG 2409  
 QY 456 GluAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val\*\*\*SerAsp----- 471  
 DB 2410 AATGCTAAAGAGACGCAATCAAGAAATTAAGAAAGCAGCAATCAGTTCTGCATATAC 2469  
 QY 472 -----LysValLysLeuVal\*\*\*AspIleLeuAlaPheLeuAlaProfile 486  
 DB 2470 TTTGATGCAATCAACAAAGCAAAACAGTAGAGCGCTAGAGCGTTTAAAGAACGAAATC 2529  
 QY 487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAsp 504  
 DB 2530 TTTAAAGCACACCTGAAACACCGCGAA----- 2559  
 QY 505 GluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAsp 524  
 DB 2559 ----- 2559  
 QY 525 ProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSer 544  
 DB 2560 -----AACCCAGGAATCACAATTGAT 2580

2Y 545 HistTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyrAla 564  
Db 2581 GAATGGTTATTAAAGTAATGCT-----AAAGAAGCTGCAATCAAGAATTA 2625  
2Y 565 LysGluLysGluLeuThrProSerThrAspHisGlnAspSerGlyAsnThrGluAla 584  
Db 2626 AAAGAAGCGAGGATCACT----- 2643  
2Y 585 LysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArg 604  
Db 2644 -----GCTGATATCTATTCACTTAATCAACAAGCAAAACAGTAGAA----- 2688  
2Y 605 MetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleProHis 624  
Db 2689 -----GGCTAGATATCTTAAAGAACGAAATCTTTAAAGACAC 2727  
2Y 625 TyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeuTyrGluAlaPro 644  
Db 2728 GCTGAATAAACCGGCGAAACCCAGGAATCACAAATTGATGATGGTTATTAAAGACGCT 2787  
2Y 645 LysGlyTyrThrLeuGluAspLeuLeu-----AlaThrValLysTyrTyrValGlu 661  
Db 2788 AAAGAAGATGCAATTAAAGATTAAGAGACGAGGATTAATCTGACATATACTTTGAT 2847  
2Y 662 HisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGln 681  
Db 2848 GCTATCAACAAA-----GCAAAAACCTATTGAAGCGGTAGAA 2883  
2Y 682 ArgAsnLysAsn-----GlyGlnAlaAspThrAsnGlnThrGluLysPro 696  
Db 2884 GCATTAAAGATGAATCTTTAAAGGCTCATATAAAGATGAAGAACAGGTATAAACCACCA 2943  
2Y 697 SerGluGlu-----LysProGlnThrGluLysPro-----GluGluGluThrProArgGluGlu 714  
Db 2944 GGTGAAGACAAATAAAGACAGATATAAACCAGGTGAAGATATAAACCAGAACACAA 3003  
2Y 715 LysPro-----GlnSerGluLysProGluSerProLysProThrGluGluProGluLysSer 733  
Db 3004 AAACCTGGTGAAGATAAATAAACCAGAGACAAATAAACCAGGTAAAGATTTCT 3063  
2Y 734 ProGluGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArgGlu 753  
Db 3064 CCAATAAAGAGAGAAAAA-----GCTAAATATCAAAAGAGCTGTAGCGAA 3108  
2Y 754 AlaGluAspLeu 757  
Db 3109 GCTGAATCTTTA 3120

RESULT 48  
JS-09-535-008-62  
Sequence 62, Application US/09535008  
Patent No. 6465629  
GENERAL INFORMATION:  
APPLICANT: Wong, Alexander K.C.  
APPLICANT: Tavtigian, Sean V.  
APPLICANT: Teng, David H.-F.  
TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE  
TITLE OF INVENTION: AND OTHER CANCER TYPES  
FILE REFERENCE: 2318-259  
CURRENT APPLICATION NUMBER: US/09/535,008  
CURRENT FILING DATE: 2000-03-23  
EARLIER APPLICATION NUMBER: U.S. 60/125,806  
EARLIER FILING DATE: 1999-03-23  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 62  
LENGTH: 5471  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (75)..(1703)  
JS-09-535-008-62

Alignment Scores:  
Pred. NO.: 0.00111 Length: 5471  
Score: 149.50 Matches: 146  
Percent Similarity: 30.35% Conservative: 81  
Best Local Similarity: 19.52% Mismatches: 241  
Query Match: 3.72% Indels: 280  
DB: 4 Gaps: 31

US-09-765-272A-66 (1-763) x US-09-535-008-62 (1-5471)

QY 177 GluAspThrGlyAspAlaTyr-----IleValProHis 187  
Db 353 GAAAGGAATGGGATGCGGTGAGGGGGCATGCTGGGATGGGGCCCCCCCCCAGCGCCCAT 412  
QY 188 Gly-AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAla 207  
Db 413 GGACGAGCACTCCCAAGGTATACCC-----TCGCCCTGGTGGCTCTGAGCATGCTCTAG 469  
QY 207 agluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerTyrAsnAl 227  
Db 470 TCCAGTT-----CCAGCCAGTGGCGGCTCTCGGGGCCCCAGATGCTTC 514  
QY 227 aAsnPro-----AlaGlnProArgLeuSerGluAsnHisAsnLe 240  
Db 515 CGGGCCAGAGGTGCGCGCTGGATGGTGTGACCCCGAGGCTTGGGGCAGCAGAACCG 574  
QY 240 uThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgG 260  
Db 575 GGGCCCAACCCCA-----TTTAAACGAGAACCG-----CTGCACCA 610  
QY 260 uLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPr 280  
Db 611 GCTCAGAGCTCAGATCATGGCTACAGATGCTG----- 644  
QY 280 oAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHi 300  
Db 645 -----GCCAGGGGGCAGCCCTCCCGCAGCACCTGCAGATGGC 682  
QY 300 sPheIleProTyrGluGlnMetSerGluLeuLysArgIleAlaArgIle----- 317  
Db 683 GGTGAGGGCAAGCGCGCATGTCGCGGATGAGCAGCAGCAGTGCACCAACGCTACCTCCACC 742  
QY 317 ----- 317  
Db 743 CTCGGTGTCCGCAACAGGACCGCGCTGGCCCTGGCCCTGGCCCGCGCGGTCCCGG 802  
QY 318 -----IleProLeuArgTyrArgSerAsnHis----- 326  
Db 803 CCGGSCACCTCCAAATTACAGCAGGCGCTCATGTATGGAGGGGCCCAACATGCTCCCCC 862  
QY 327 -----TrpValProAspSerArgProGluGlnPro----- 336  
Db 863 AGGACCTCGGGCGTGGCCCGGGATGCCAGGCCAGCCTCTCGAGGGCCCTCCCAAGCC 922  
QY 337 -----SerProGlnSerThrProGlu----- 343  
Db 923 CTGGCTGAAGACCCATGGCGAATGTGTGCTGCCCCACAGCAGCACCTTCAGAGAGTGTAT 982  
QY 344 -----ProSerProSerProGlnProAlaProAsnPro----- 354  
Db 983 TCCCCCGCAGCAACGGCGCGGCTTCCCGCGCGCCCTGCGTCCCGCGCGCTC 1042  
QY 355 -----GlnProAlaProSerAs 360  
Db 1043 GCCCGTGTATGCCACCGCAGACCCAGTCCCGCGGCGAGCGCGCGCGCCCATGTGT 1102  
QY 360 nProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGly----- 377  
Db 1103 GCCACTGCACCAAGACGCGCGCTCATCCCGCATCCAGAGACCGCGGGGCTCCAGCCC 1162  
QY 378 -----TyrValPheGluGluAsnGlyValSerArgTyrI 389

Db 1163 TGTGAGATCTCTGAGGAGCGGAGTACAGGCTGCAGGCTCGCATCGCACACCGGAATCA 1222  
QY 389 eProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerIlys----- 405  
Db 1223 GGAACCTTTGAAACCTTCCCGGGTCCCTGCGGGGATTTGGCAACCAAGGACCATGGA 1282  
QY 405 ----- 405  
Db 1283 GCTCAAGCCCTCAGGCTGTGAACCTCCAGGAGCGCTGCGCCAGGAGGTGGTGGTGG 1342  
QY 406 -LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys-----LysThrAspLe 423  
Db 1343 CATCGGAGGACACAGCGCTGGAGACAGCCCTCAATCTTAAGGCTACAGCGGAGCAAA 1402  
QY 423 uProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisG 443  
Db 1403 GCGCCAGTCCCTGCGGAG-----GCCCGCATCACTGA 1435  
QY 443 nAspLeu-----LeuAspAsnLysGlyArgGlnValAspPheG 456  
Db 1436 GAAGCTGGAGAGCAGCAGAGATCGAGCAGGCGCAAGCGCCGCGCAGACACCGAGA 1495  
QY 456 uAlaLeuAspAsnLeuLeuGluArgLeuLysAsp-----Val\*\* 469  
Db 1496 ATACCTCAATAGCATCTCCAGCATGCCAAGATTCAGGAATATCACAGATCGGTAC 1555  
QY 469 \*SerAspLysValLysLeuVal\*\*\*AspIleLeuAlaPheLeuAlaProIle----- 486  
Db 1556 AGGCAAAATCCAGAGCTGACCAAGGAGTGCACGATACCATGCCAACACAGGCGGGA 1615  
QY 487 -----ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAs 504  
Db 1616 GCAGAGAAAGAGAACAGCGGATCGAGAGAGCGATGCGGAGGCTCATGGCTGAGA 1675  
QY 504 pGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAs 524  
Db 1676 TGAGGAGGTACCGCAAGCTCATC----- 1700  
QY 524 pProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisE 544  
Db 1701 -----GACTAGAAAGAGCAAGCGCTGCTTACTCTTG--CAGCAGACAGACGA 1750  
QY 544 rHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAl 564  
Db 1751 GTACGTGGCTTAACCTCACGAGCTGTGTGCGCAGCACAGAGCTGCCAGGTCGCCAAGA 1810  
QY 564 aLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAl 584  
Db 1811 GAAAGAGAGAAAGAAAGAAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1867  
QY 584 aLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspAr 604  
Db 1868 CATGGGCGGATGGC-----GAGCCTCTAGACGA 1897  
QY 604 g-----MetProTyrAsnLeuGlnTyrThrValGluValLysAsnG 618  
Db 1898 GACCAGCCAGATGAGCGACCTCCCGGTGAAGGTG-----ATCCAGCTGAGAGGTG 1948  
QY 618 ySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspG 638  
Db 1949 GAAGATCTCCACAGGC----- 1964  
QY 638 uGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTy 658  
Db 1965 -----ACAGATGCCCCCAAGCCCGGAGCTGGAGCTGGCTCAGATGAACCCGGG 2017  
QY 658 rTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAs 678  
Db 2018 GTAT-----GAAGTAGCTCCGAGGTCTGAT----- 2042  
QY 678 pHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerG 698  
Db 2043 -----AGTGA 2047

QY 698 uGluLysProGlnThrGluLysProGluGluThrProArgGluGluLysProGlnSe 718  
Db 2048 AGAAAGTGTCTCAGAAAG 2101  
QY 718 rGluLysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerG 738  
Db 2102 AGCAGACGCTCCACCC-----CTGCCCTGGAGGAGAGAGAGAGATTCAGCATCCAGACAG 2158  
QY 738 uGluProGlnValGluThrGluLys-----ValGluGluLysLeuArgGluAlaG 755  
Db 2159 CGATGACGCTCTGAGTGGAGCGCGGGGCACATCATTTGAGATGCCAAGCAAGATGTGCA 2218  
QY 755 uAspLeuLeuGlyLysIleGln 762  
Db 2219 TGATGAATATGCGGTGTCCAG 2240

## RESULT 49

US-09-103-840A-2

; Sequence 2, Application US/09103840A

; Patent No. 6294328

; GENERAL INFORMATION:

; APPLICANT: FLEISCHMAN, Robert D.

; APPLICANT: WHITE, Owen R.

; APPLICANT: FRASER, Claire M.

; APPLICANT: VENTER, John C.

; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM

; FILE REFERENCE: 24366-20007.00

; CURRENT APPLICATION NUMBER: US/09/103,840A

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 4403765

; TYPE: DNA

; ORGANISM: Mycobacterium tuberculosis

; FEATURE:

; OTHER INFORMATION: CDC 1551

; OTHER INFORMATION: "n" bases at various positions throughout the sequence

; OTHER INFORMATION: represent a, t, c or g

US-09-103-840A-2

## Alignment Scores:

Pred. No.:	48.7	Length:	4403765
Score:	149.50	Matches:	161
Percent Similarity:	30.71%	Conservative:	89
Best Local Similarity:	19.78%	Mismatches:	284
Query Match:	3.72%	Indels:	282
DB:	3	Gaps:	37

US-09-765-272A-66 (1-763) x US-09-103-840A-2 (1-4403765)

QY 129 ArgThrLysGluGluLysArgGlnLysGlnGluArgSerHisAsnHisAsnSer--- 147

Db 2479156 AGGCCAGAGTGGAGCTGCGCGGACCGACGATCGAGGAGTCAACAGAAATGGCCTT 2479215

QY 148 ---ArgAlaAspAsnAlaValAlaAlaArgAlaGlnGlyArgTyrThrThrAspAsp 166

Db 2479216 CTCGCTCCAGATGCGGACCTCGGTGAGAGCGCTCCACGAGG---GAC 2479260

QY 167 GlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleValPro 186

Db 2479261 GGTAC-----CCG 2479269

QY 187 HisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAla 206

Db 2479270 CT-GGCTCAACAGGAGGCGACACGGTCACTCGACGAGCCCTCG----- 2479316

QY 207 AlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerSerTyrAsn 226

Db 2479317 -----TGAGG-----TGTCGACCGCAAGGTCGACACCG----- 2479346

[illegible]

RESULT 50

US-09-103-840A-1

```

; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 48.8 Length: 4411529
Score: 149.50 Matches: 161
Percent Similarity: 30.71% Conservative: 89
Best Local Similarity: 19.78% Mismatches: 284
Query Match: 3.72% Indels: 282
DB: 3 Gaps: 37

US-09-765-272A-66 (1-763) x US-09-103-840A-1 (1-4411529)
QY 129 ArgThrLysGluGluLeuLysArgGlnLysGlnGluArgSerHisAsnHisAsnSer--- 147
Db 2481911 AGGCGAGAGTGGAGCTCGCGCGACCGACGATCGATCGAGAGTCAACAGAAATGGCCTT 2481970
QY 148 ---ArgAlaAspAsnAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 166
Db 2481971 CTCGCTCCAGATGCCGGCAGCTCGGTGAGAGCGTCAACGAGG---GAC 2482015
QY 167 GlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleValPro 186
Db 2482016 GGTTC---CGG 2482024
QY 187 HisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAla 206
Db 2482025 CT-GGCTCAACAGAGAGGCGACACGGTGAACCTCGACGAGCCCTCG--- 2482071
QY 207 AlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSerSerSerSerTyrAsn 226
Db 2482072 ---TGGAGG---TGTCACCGCAAGAGTTCGACACCG--- 2482101
QY 227 AlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThrTyr 246
Db 2482102 AAATCCCTCGCGCGCGCGGTGTGTGACCAAGATCATCGCCAGGAGGATGACGG 2482161
QY 247 HisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysProLeu 266
Db 2482162 TCGAGGTGGCGGCGAGTCTGCTGTCTATGTGGCGAGCGCAAGAGTTCGCGGAGGCGG--- 2482218
QY 267 SerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArg 286
Db 2482219 CGGCGCGCGGACCCGAGAAAGTCC---CTGCGCGCGCAACCCGAGTCCAGCGCGG 2482269
QY 287 ThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGluGln 306
Db 2482270 ---CACCGAACACAC---CGG 2482287
QY 307 MetSerGluLeuGluLysArgIleAlaAlaGlnIlePro--- 319
Db 2482288 TCCAAACGACGCTCCGAGCGCTGTGTGTGGCGATGCCAAGCCGCTGTGTGTCGCCGAGC 2482347
QY 320 ---LeuArg 321
Db 2483322 AAAACC-----GACGAGTTGTCCGCGGCGACCTTCACCATCACCACATCGGTAGC 2483372

```

```

Db 2483348 TCGCGCGAATCGGTACCGAGGGGACCGTCACTTCGTTGGCTGAAGAAGATCGGGATTCCG 2483407
QY 322 TyrArg---SerAsnHis-----TyrValProAspSerArg 332
Db 2483408 TTCAGGTTGACGAGCCACTCGTGGAGGTGCCACCGACAAAGTGGACACCGAGATCCCGT 2483467
QY 333 Pro----- 333
Db 2483468 CCCCGGTGGTGGGGTCTTGGTCAGTATCATCGCGCCGACGAGGACGCCACGGTGGCCGTCG 2483527
QY 334 -----GluGlnProSerProGlnSerThr 341
Db 2483528 GCGCGAGTTGGCCCGGATCGGTCTCGTCCGACATCGCGCGCGCGCCGCCCAAGC 2483587
QY 342 ProGluProSerProSerProGlnProAlaProAsnPro-----GlnProAlaProSer 359
Db 2483588 CCGCACCCCAAGCCCGTCCCGAG---CCAGCGCGGAGCGCGAAGCGCGACCCGACCATCG 2483646
QY 360 AsnProIleAspGlu-----LysLeuValLysGluAlaVal 371
Db 2483647 CCGCGCGCGCGCCAGCCAGCCGCGTGGCGCGAGGCGCACCGTACGTGACGCCGTGGTG 2483706
QY 372 ArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSerArgTyrIleProAla 391
Db 2483707 CGAAAGCTGGCG-----TCGAAACCAACATC----- 2483733
QY 392 LysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSer 411
Db 2483734 ---GACCTCGCGGGGTGACCGGCAACGAGTGGTGGTTCGATCCGCAACACAGATGTG 2483790
QY 412 LeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSerSerAsp--- 427
Db 2483791 CTGGCGCGCGCTGAGCAAAAGAGCGGCAACGACCGCGCGCGCGCGCGCGCGCGCC 2483850
QY 428 -----ArgGlu 429
Db 2483851 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2483910
QY 430 PheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeuAspAsnLys 449
Db 2483911 ACCACCCAGAGGCC-----AGCCGATTCGTGAG---ATCACCGCCACACAG 2483955
QY 450 GlyArg-----GlnValAspPheGlu 456
Db 2483956 ACCCGGAATCTTTGACGCAACGCGCACAGCTGACACAAACCCATGAGTTCGATGACC 2483015
QY 457 AlaLeuAspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal 476
Db 2483016 AAGATCGTGGGGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2483075
QY 477 ***AspIleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsn 496
Db 2483076 CTGACCTTCTGCGGTCTTCGCCAAGCGCGTGTATGATGCCCTCAAGATTCAACCGAAC 2483135
QY 497 AlaGlnIleThrTyrThrAspAspGluIleGlnValAla-----LysLeu 511
Db 2483136 ATCAACGCTAGCTACACGAGGACACCAAGAGATCACCTACTACGACGCGCGGACCTA 2483195
QY 512 AlaGlyLysTyrThrThrGluAspGlyTyrIlePheAspProArgAspIleThrSerAsp 531
Db 2483196 GGATTCTGCTGCGACACCGGAGCGCGCTGCTC-----TCCCGGTTCATCCACGAC 2483246
QY 532 GluGlyAspAlaTyrValThrProHisMetThrHisSerHisTrpIleLysLysAspSer 551
Db 2483247 GCCGCGATCTG-----TCA 2483261
QY 552 LeuSerGluAlaGluArgAlaAlaAlaGln-----AlaTyrAlaLysGluLysGlyLeu 569
Db 2483262 CTGGCGGTCTGGCGGCGGCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2483321
QY 570 ThrProProSerThrAspHisGlnAspSerGly-----AsnThrGluAla 584
Db 2483322 AAAACC-----GACGAGTTGTCCGCGGCGACCTTCACCATCACCACATCGGTAGC 2483372

```

595	LysGlyAla-	:::	-----GluAlaIleTyrAsnArgVal	594
2483373	CAGGCGCGTTGTTCGACACC	CCCGATCTGGTTCGCCG	CAGGCAGCCCATCTGTGGCACC	2483432
595	LysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGlu	614		
2483433	GGGCGGATCGTCAACGCGCGCGGGTGCTCGATGCCAGCGCAACAGTTCGATCGGG	2483492		
615	VallYsAsnGlySerLeuIleIleProHisTyrAspHis-	627		
2483493	GTCGCTCGTCTGTACTCTCCGTTGACCTATGACCATCGCGCTCATCGACGCGCGCAC	2483552		
628	-----TyrHisAsnIleLysPheGluTrpPheAspGluGlyLeuTyrGluAla	643		
2483553	GCCGCGCTTTCCTCACCACGATCAAGCACCGC--CTCGAAGAGGAGCGTTCGAGGCC	2483609		
644	ProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTrpVal	660		
2483610	GATTTAGCACTGTGATGGTGCACACGCGTGTGCGCATCGCGGTTCGTCTGCGTTGATCG	2483669		
661	-----GluHisProAsnGluArgProHisSerAspAngLy	672		
2483670	GCTCTGCGCTGACCGCGCGCTGTCGCGCGCGCGACACACCG--	TGC	2483714	
673	PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsn-	-GlyGln	687	
2483715	TGCGGATCGTGGCGCGGACCTGCGAATTCGGAAGACTGCATGGATCCGAAAGCG	2483774		
688	AlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLysProGlu	707		
2483775	GCGAATTCGATCCGACGCGCTCACCGATGTCGACCGCGTGTCAACCTCTCGCGGCTCA	2483834		
708	GluGlu-	-----ThrProArgGlu	713	
2483835	ACATGCCCGACGTCGGTGTGCGGGCTTCAAACAGAGCCTGCGCAGACCGCGATCA	2483894		
714	GluLysProGlnSerGluLysProGluSerProLysProThrGluGluProGluGluSer	733		
2483895	CACCCACGAGGTGCTATCCGCGCGAGTCGCCAGCGCGCGTCTGCTACCTTATCAACG	2483954		
734	ProGluGlu-	-----SerGluGluProGln-	ValGluThrGluLys	746
2483955	CCAGCGCGTGGCTACTACGGAACAACAAGGACCGGGTGTGTCGACGAAACGATCTCG	2484014		
747	ValGluGluLysLeuArgGluAlaGluAspLeuGlyLys	760		
2484015	CGGGAACAGGTTTTCTGGCCCCAGCTGTGCTTGA	2484056		



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

DM nucleic - nucleic search, using sw model

Run on: May 1, 2004, 02:02:14 ; Search time 163 Seconds  
(without alignments)  
7796.549 Million cell updates/sec

Title: US-09-765-272A-65

Perfect score: 2290

Sequence: 1 TTGTCCTATGAACCTGTC.....TACTTGGAAATCCAGGAT 2290

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

Issued Patents\_NA.\*

1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq\*

2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq\*

3: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq\*

4: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq\*

5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2285	99.8	2290	3	US-08-961-083-65	Sequence 65, Appl
2	2285	99.8	2290	4	US-09-536-784-65	Sequence 65, Appl
3	2220.2	97.0	2531	4	US-09-468-656A-11	Sequence 11, Appl
4	1999.6	87.3	2531	4	US-09-468-656A-5	Sequence 5, Appl
5	1999.2	87.3	8195	4	US-08-961-527-94	Sequence 94, Appl
6	1376.6	60.1	2359	4	US-08-961-527-243	Sequence 243, App
7	990.6	43.3	2451	4	US-09-468-656A-9	Sequence 9, Appl
8	987.6	43.1	2389	3	US-08-961-083-55	Sequence 55, Appl
9	987.6	43.1	2389	4	US-09-536-784-55	Sequence 55, Appl
10	794.8	34.7	973	4	US-08-961-527-355	Sequence 355, App
11	677	29.6	1684	4	US-08-961-527-258	Sequence 258, App
12	401	17.5	1342	3	US-08-961-083-181	Sequence 181, App
13	401	17.5	1342	4	US-09-536-784-181	Sequence 181, App
14	401	17.5	1455	4	US-09-468-656A-7	Sequence 7, Appl
15	401	17.5	6867	4	US-08-961-527-192	Sequence 192, App
16	272.4	11.9	841	2	US-08-743-637B-34	Sequence 34, Appl
17	272.4	11.9	841	3	US-08-526-840B-34	Sequence 34, Appl
18	64.4	2.8	7218	1	US-08-232-463-14	Sequence 14, Appl
19	56.2	2.5	2223	1	US-08-257-073-4	Sequence 4, Appl
20	51.4	2.2	865	4	US-09-640-419C-6	Sequence 6, Appl
21	47.8	2.1	31096	4	US-08-956-171E-59	Sequence 59, Appl
22	46.6	2.0	2787	4	US-09-134-000C-3185	Sequence 3185, App
23	45.4	2.0	5340	4	US-09-627-122-21	Sequence 21, Appl
24	44.8	2.0	3337	1	US-08-072-610-1	Sequence 1, Appl
25	44.8	2.0	3337	2	US-08-719-822B-1	Sequence 1, Appl
26	44.8	2.0	3337	3	US-09-092-458-1	Sequence 1, Appl
27	44.2	1.9	43795	3	US-08-742-185-101	Sequence 101, App

28	43.8	1.9	765	4	US-08-914-479A-3	Sequence 3, Appl
29	43.8	1.9	855	4	US-08-914-479A-5	Sequence 5, Appl
30	43.8	1.9	918	3	US-08-937-271-9	Sequence 9, Appl
31	43.2	1.9	3095	6	5231168-1	Patent No. 5231168
32	43	1.9	478	4	US-09-621-976-10407	Sequence 10407, A
33	42.6	1.9	1284	4	US-09-286-981B-24	Sequence 24, Appl
34	42.4	1.9	929	4	US-03-671-317-14	Sequence 14, Appl
35	42.4	1.9	1236	4	US-09-286-981B-34	Sequence 34, Appl
36	42.4	1.9	1338	4	US-09-286-981B-23	Sequence 23, Appl
37	42.4	1.9	5295	4	US-09-221-017B-27	Sequence 27, Appl
38	42.2	1.8	2766	4	US-09-543-681A-1562	Sequence 1562, Ap
39	42	1.8	1239	4	US-09-286-981B-22	Sequence 22, Appl
40	42	1.8	1275	4	US-09-286-981B-28	Sequence 28, Appl
41	41.4	1.8	696	3	US-09-461-697-193	Sequence 193, App
42	41.4	1.8	699	3	US-09-461-697-191	Sequence 191, App
43	41.4	1.8	717	3	US-09-461-697-189	Sequence 189, App
44	41.4	1.8	774	3	US-09-461-697-187	Sequence 187, App
45	41.4	1.8	819	3	US-09-461-697-185	Sequence 185, App
46	41.4	1.8	1001	4	US-09-671-317-439	Sequence 439, App
47	41.4	1.8	1360	3	US-08-961-083-37	Sequence 37, Appl
48	41.4	1.8	1360	4	US-09-536-784-37	Sequence 37, Appl
49	41.4	1.8	1669	3	US-09-461-697-184	Sequence 184, App
50	41.4	1.8	12665	4	US-08-961-527-134	Sequence 134, App
51	41	1.8	1218	4	US-08-286-981B-35	Sequence 35, Appl
52	40.8	1.8	1338	4	US-09-286-981B-26	Sequence 26, Appl
53	40.8	1.8	1830121	4	US-09-557-884-1	Sequence 1, Appl
54	40.8	1.8	1830121	4	US-09-643-990A-1	Sequence 1, Appl
55	40.4	1.8	1280	4	US-09-220-132-155	Sequence 155, App
56	40.2	1.8	1815	2	US-08-687-956A-2	Sequence 2, Appl
57	40	1.7	775	3	US-08-961-083-89	Sequence 89, Appl
58	40	1.7	775	4	US-09-536-784-89	Sequence 89, Appl
59	40	1.7	1696	3	US-08-961-083-217	Sequence 217, App
60	40	1.7	1696	4	US-09-536-784-217	Sequence 217, App
61	40	1.7	3211	2	US-08-574-959A-8	Sequence 8, Appl
62	40	1.7	3211	3	US-08-357-014-8	Sequence 8, Appl
63	40	1.7	3901	2	US-08-574-959A-6	Sequence 6, Appl
64	40	1.7	3901	3	US-09-357-014-6	Sequence 6, Appl
65	40	1.7	13926	4	US-08-961-527-5	Sequence 5, Appl
66	40	1.7	1830121	4	US-09-557-884-1	Sequence 1, Appl
67	40	1.7	1830121	4	US-09-643-990A-1	Sequence 1, Appl
68	39.8	1.7	4488	4	US-08-956-171E-228	Sequence 228, App
69	39.8	1.7	4858	4	US-08-956-171E-379	Sequence 379, App
70	39.6	1.7	1236	2	US-08-741-134-5	Sequence 5, Appl
71	39.6	1.7	2070	4	US-09-134-000C-1094	Sequence 1094, Ap
72	39.6	1.7	1664976	4	US-08-916-421B-1	Sequence 1, Appl
73	39.4	1.7	395	4	US-09-894-844-45	Sequence 45, Appl
74	39.4	1.7	1258	4	US-09-286-981B-32	Sequence 32, Appl
75	39.4	1.7	2222	4	US-08-956-171E-380	Sequence 380, App
76	39.4	1.7	6755	3	US-08-931-999-4	Sequence 4, Appl
77	39.4	1.7	11091	4	US-09-134-001C-2243	Sequence 2243, Ap
78	39.2	1.7	3466	1	US-08-468-036-38	Sequence 38, Appl
79	39.2	1.7	3466	2	US-08-376-843-38	Sequence 38, Appl
80	39.2	1.7	5552	3	US-08-155-888-1	Sequence 1, Appl
81	39	1.7	8009	4	US-08-956-171E-61	Sequence 61, Appl
82	38.8	1.7	396	4	US-09-640-173-53	Sequence 53, Appl
83	38.8	1.7	396	4	US-09-713-550-53	Sequence 53, Appl
84	38.8	1.7	2462	3	US-08-847-065-24	Sequence 24, Appl
85	38.8	1.7	2793	2	US-08-718-661-1	Sequence 1, Appl
86	38.8	1.7	3023	3	US-09-308-022-4	Sequence 4, Appl
87	38.8	1.7	9636	1	US-08-323-170B-1	Sequence 1, Appl
88	38.8	1.7	9636	1	US-08-954-441-1	Sequence 1, Appl
89	38.6	1.7	376	2	US-08-623-906A-18	Sequence 18, Appl
90	38.6	1.7	1332	2	US-08-795-475-5	Sequence 5, Appl
91	38.6	1.7	2133	4	US-08-107-532A-1413	Sequence 1413, Ap
92	38.6	1.7	3837	4	US-09-724-517-1	Sequence 1, Appl
93	38.6	1.7	3837	4	US-09-641-807A-1	Sequence 1, Appl
94	38.6	1.7	3837	4	US-09-723-096-1	Sequence 1, Appl
95	38.4	1.7	754	4	US-09-286-981B-21	Sequence 21, Appl
96	38.4	1.7	1242	4	US-09-286-981B-27	Sequence 27, Appl
97	38.4	1.7	1242	4	US-09-286-981B-33	Sequence 33, Appl
98	38.4	1.7	5361	3	US-08-973-462-2	Sequence 2, Appl
99	38.4	1.7	6152	3	US-08-973-462-1	Sequence 1, Appl
100	38.2	1.7	289	3	US-09-007-005-17	Sequence 17, Appl

101 38.2 1.7 289 3 US-09-244-796-17 Sequence 17, Appl  
102 38.2 1.7 966 2 US-08-766-738-2 Sequence 2, Appl1  
103 38.2 1.7 966 2 US-09-262-610-2 Sequence 2, Appl1  
104 38.2 1.7 3006 4 US-09-762-724-3 Sequence 3, Appl1  
105 38.2 1.7 5394 3 US-08-688-376-1 Sequence 1, Appl1  
106 37.8 1.7 387 4 US-09-621-976-8572 Sequence 8572, Ap  
107 37.6 1.6 2169 4 US-09-434-408-3 Sequence 3, Appl1  
108 37.6 1.6 2447 2 US-09-014-969-14 Sequence 14, Appl  
109 37.6 1.6 3135 4 US-09-107-532A-1575 Sequence 1575, Ap  
110 37.6 1.6 4322 2 US-08-537-342-1 Sequence 1, Appl1  
111 37.6 1.6 10640 4 US-09-417-485D-5 Sequence 5, Appl1  
112 37.4 1.6 6617 4 US-09-689-065B-1 Sequence 1, Appl1  
113 37.2 1.6 658 4 US-09-286-981B-25 Sequence 25, Appl  
114 37 1.6 1203 4 US-09-134-001C-1483 Sequence 1483, Ap  
115 37 1.6 1695 3 US-09-308-022-5 Sequence 5, Appl1  
116 37 1.6 2780 4 US-09-220-132-176 Sequence 176, Ap  
117 37 1.6 2780 4 US-09-220-132-178 Sequence 178, Ap  
118 36.8 1.6 4081 3 US-08-999-774A-1 Sequence 1, Appl1  
119 36.6 1.6 240 1 US-08-628-417-6 Sequence 6, Appl1  
120 36.6 1.6 657 4 US-09-527-345-3 Sequence 3, Appl1  
121 36.6 1.6 2167 4 US-09-216-393B-326 Sequence 326, Ap  
122 36.6 1.6 2167 4 US-09-216-393B-328 Sequence 328, Ap  
123 36.6 1.6 2392 4 US-09-216-393B-331 Sequence 331, Ap  
124 36.6 1.6 2393 4 US-09-216-393B-329 Sequence 329, Ap  
125 36.6 1.6 5562 4 US-10-204-708-63 Sequence 63, Appl  
126 36.4 1.6 888 4 US-09-134-001C-1629 Sequence 1629, Ap  
127 36.4 1.6 956 4 US-08-714-741-36 Sequence 36, Appl  
128 36.4 1.6 1056 4 US-09-328-352-2030 Sequence 2030, Ap  
129 36.4 1.6 1990 4 US-08-714-741-35 Sequence 35, Appl  
130 36.4 1.6 2085 1 US-08-072-070-1 Sequence 1, Appl1  
131 36.4 1.6 2085 1 US-08-465-746-1 Sequence 1, Appl1  
132 36.4 1.6 2085 1 US-08-214-164-1 Sequence 1, Appl1  
133 36.4 1.6 2085 1 US-08-469-434-1 Sequence 1, Appl1  
134 36.4 1.6 2085 1 US-08-214-222-1 Sequence 1, Appl1  
135 36.4 1.6 2085 2 US-08-467-852A-1 Sequence 1, Appl1  
136 36.4 1.6 2085 2 US-08-468-718-1 Sequence 1, Appl1  
137 36.4 1.6 2085 2 US-08-246-636-1 Sequence 1, Appl1  
138 36.4 1.6 2085 2 US-08-247-491A-1 Sequence 1, Appl1  
139 36.4 1.6 2085 2 US-08-319-795-1 Sequence 1, Appl1  
140 36.4 1.6 2085 2 US-08-468-985-1 Sequence 1, Appl1  
141 36.4 1.6 2086 3 US-08-312-949-1 Sequence 1, Appl1  
142 36.4 1.6 2086 3 US-08-446-201-2 Sequence 2, Appl1  
143 36.4 1.6 4766 5 PCT-US93-07261-10 Sequence 10, Appl  
144 36.2 1.6 429 3 US-08-847-065-20 Sequence 20, Appl  
145 36.2 1.6 1071 4 US-09-134-000C-3347 Sequence 3347, Ap  
146 36.2 1.6 10993 4 US-08-961-527-15 Sequence 15, Appl  
147 36.2 1.6 72604 4 US-09-268-992-7 Sequence 7, Appl1  
148 36.2 1.6 72604 4 US-09-657-474-7 Sequence 7, Appl1  
149 36 1.6 1436 4 US-09-023-655-873 Sequence 873, Ap  
150 36 1.6 1453 4 US-08-714-741-33 Sequence 33, Appl

## ALIGNMENTS

RESULT 1  
US-08-961-083-65  
; Sequence 65, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: Hp Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,083  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2290 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-961-083-65  
  
Query Match 99.8%; Score 2285; DB 3; Length 2290;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TTCTTCTATGAAGTGGTCTGCTCACCAGCTGGTTCAGGTTAAGGAGGATCTAATCGAGT 60  
DB 1 TTCTTCTATGAAGTGGTCTGCTCACCAGCTGGTTCAGGTTAAGGAGGATCTAATCGAGT 60  
  
QY 61 TTCTTATATAGATGGTGCATCAGGCTGGTCAAAAGGCGAGAAAATTGACACAGATGAAGT 120  
DB 61 TTCTTATATAGATGGTGCATCAGGCTGGTCAAAAGGCGAGAAAATTGACACAGATGAAGT 120  
  
QY 121 CAGTAAGAGGAGGGGATCAACCGCGAAATATGTTATCAAGATTAACGATCAAGGTTA 180  
DB 121 CAGTAAGAGGAGGGGATCAACCGCGAAATATGTTATCAAGATTAACGATCAAGGTTA 180  
  
QY 181 TGTGACCTCTCATGGAGACCAATATCAATTAATTAATGGCAAGGTTCTTATGATGCCAT 240  
DB 181 TGTGACCTCTCATGGAGACCAATATCAATTAATTAATGGCAAGGTTCTTATGATGCCAT 240  
  
QY 241 CATCAGTGAAGAGCTCTCATGAAGATCCGAATTCAGTTGAAGATTCAGACATTTGT 300  
DB 241 CATCAGTGAAGAGCTCTCATGAAGATCCGAATTCAGTTGAAGATTCAGACATTTGT 300  
  
QY 301 CAATGAAATCAAGGGTGGTTATGTCAATTAAGGTTAAACGGTAAATCTATGTTACCTTAA 360  
DB 301 CAATGAAATCAAGGGTGGTTATGTCAATTAAGGTTAAACGGTAAATCTATGTTACCTTAA 360  
  
QY 361 GGATGAGCTCATGGGATTAATTCGGACAAAAGAGAGATTAAACGTCAAGACGAGA 420  
DB 361 GGATGAGCTCATGGGATTAATTCGGACAAAAGAGAGATTAAACGTCAAGACGAGA 420  
  
QY 421 ACCGAGTCAATATCATCAATCAAGAGCAGATAATGCTGTTGCTCAGCCAGGCCAAGG 480  
DB 421 ACCGAGTCAATATCATCAATCAAGAGCAGATAATGCTGTTGCTCAGCCAGGCCAAGG 480  
  
QY 481 ACCTTATCAACCGGATGATGGGTATATCTTCAATGATCTGATATCAATTTAGGACACGG 540  
DB 481 ACCTTATCAACCGGATGATGGGTATATCTTCAATGATCTGATATCAATTTAGGACACGG 540  
  
QY 541 TGATGCTTATATCGTTCCTCAACGGGACCAATTAACATTAATTCCTTAAGATGATGTTATC 600  
DB 541 TGATGCTTATATCGTTCCTCAACGGGACCAATTAACATTAATTCCTTAAGATGATGTTATC 600  
  
QY 601 AGCTAGGAGTTAGTCTCGAGAGCAGATTTGGAATGGGAGCAGGATCTCGTCCCTTC 660  
DB 601 AGCTAGGAGTTAGTCTCGAGAGCAGATTTGGAATGGGAGCAGGATCTCGTCCCTTC 660  
  
QY 661 TTCAAGTTCTAGTTATAATGCAAAATCCAGCTCAACCAAGATTGTGAGAGAACCAATCT 720

661 TTCAAGTCTAGTTAAATGCAATCCAGCTCAACCAAGATTGTGAGAAACACCAATCT 720  
721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAACATTTCAAGCCTTTTACGTGA 780  
721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAACATTTCAAGCCTTTTACGTGA 780  
781 ATTGTATGTCTAAACCTTATCAGAAAGCCATGTGGAACTGTGATGGCCCTTATTTTCGACCC 840  
781 ATTGTATGTCTAAACCTTATCAGAAAGCCATGTGGAACTGTGATGGCCCTTATTTTCGACCC 840  
841 AGCGCAAAATCACAAGTCGAACCCGACAGGTGTAGCTGCTCCCTCATCTGTAACCAATACCA 900  
841 AGCGCAAAATCACAAGTCGAACCCGACAGGTGTAGCTGCTCCCTCATCTGTAACCAATACCA 900  
901 CTTTATCCCTTATGAACAAATGTCTGAATTTGGAATAACGAATTTGCTCGTATTTTCCCTT 960  
901 CTTTATCCCTTATGAACAAATGTCTGAATTTGGAATAACGAATTTGCTCGTATTTTCCCTT 960  
961 TCGTTATCGTTCAAACTTGGGTACAGATTCAGAACTCAAGACCAAGCAACCAAGTCCCAATC 1020  
961 TCGTTATCGTTCAAACTTGGGTACAGATTCAGAACTCAAGACCAAGCAACCAAGTCCCAATC 1020  
1021 GACTCCGGAACCTAGTCCCAAGTCCGCAACCTGCACCAAAATCCTCAACAGCTCCCAAGCAA 1080  
1021 GACTCCGGAACCTAGTCCCAAGTCCGCAACCTGCACCAAAATCCTCAACAGCTCCCAAGCAA 1080  
1081 TCCAAATGTATGAGAAATTTGCTCAAGAAAGCTGTTGCAAAAAGTAGGCGATGTTATGTCCTT 1140  
1081 TCCAAATGTATGAGAAATTTGCTCAAGAAAGCTGTTGCAAAAAGTAGGCGATGTTATGTCCTT 1140  
1141 TGAGGAATATGGAATTTCTGTTATATCCAGCAAGGATCTTTTCAGCAAGAAACAGCAGC 1200  
1141 TGAGGAATATGGAATTTCTGTTATATCCAGCAAGGATCTTTTCAGCAAGAAACAGCAGC 1200  
1201 AGGCATTTAGTCAAACTGCGCAAGCAGCAAGATTTATCTCATAAAGCTAGGAGCTAAGAA 1260  
1201 AGGCATTTAGTCAAACTGCGCAAGCAGCAAGATTTATCTCATAAAGCTAGGAGCTAAGAA 1260  
1261 AACTGACCTCCCATCTAGTATGATCGAATTTTACAATAAGCTTATGACTTACTAGCAAG 1320  
1261 AACTGACCTCCCATCTAGTATGATCGAATTTTACAATAAGCTTATGACTTACTAGCAAG 1320  
1321 AATTCAACCAAGATTTACTTGATATAAAGCTGCACAAAGTTGATTTTTCAGGCTTTGATAA 1380  
1321 AATTCAACCAAGATTTACTTGATATAAAGCTGCACAAAGTTGATTTTTCAGGCTTTGATAA 1380  
1381 CTTGTTGGAACGACTCAAGATGTCNCAAGTATATAAGTCAAGTTAGTGANGATATCTT 1440  
1381 CTTGTTGGAACGACTCAAGATGTCNCAAGTATATAAGTCAAGTTAGTGANGATATCTT 1440  
1441 TGCCTTTCTTAGCTCCGATTCCTCATCCAGAACTTTTAGGAAACCAATGCGCAAAATTC 1500  
1441 TGCCTTTCTTAGCTCCGATTCCTCATCCAGAACTTTTAGGAAACCAATGCGCAAAATTC 1500  
1501 CTACACTGATGATGAGATTCAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAGACGG 1560  
1501 CTACACTGATGATGAGATTCAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAGACGG 1560  
1561 TTATATCTTTGATCTCGTGATATAACCAAGTATGAGGGGATGCTTATGTAATCCACA 1620  
1561 TTATATCTTTGATCTCGTGATATAACCAAGTATGAGGGGATGCTTATGTAATCCACA 1620  
1621 TATGACCCATAGCCACTGGATTTAAAAGATGTTGCTGTAAGCTCAGAGCGGCGAGC 1680  
1621 TATGACCCATAGCCACTGGATTTAAAAGATGTTGCTGTAAGCTCAGAGCGGCGAGC 1680  
1681 CCAGGCTTATGCTTAAAGAGAAAGTTTGACCCCTCTTTCGACAGACCATCAGGATTCAGG 1740  
1681 CCAGGCTTATGCTTAAAGAGAAAGTTTGACCCCTCTTTCGACAGACCATCAGGATTCAGG 1740  
1741 AAATCTAGGCAAAAGGAGCAGAACTATCTACAACCGCGTGAAGCAGCTAAGAGGT 1800

Db 1741 AAATCTAGGCAAAAGGAGCAGAAAGCTATCTACAACCCCGTGAAGCAGCTAAGAGGT 1800  
Qy 1801 GCCACTTGATCGTATGCTTACAATCTTCAATATACTGTAGAAAGTCAAAAACGCTAGTTT 1860  
Db 1801 GCCACTTGATCGTATGCTTACAATCTTCAATATACTGTAGAAAGTCAAAAACGCTAGTTT 1860  
Qy 1861 AATCATACCTCATTTATGACCAATACCAATAACATCAAAATTTGAGTGGTGTTCGAAAGGCT 1920  
Db 1861 AATCATACCTCATTTATGACCAATACCAATAACATCAAAATTTGAGTGGTGTTCGAAAGGCT 1920  
Qy 1921 TTATGAGGCACTAAGGGGTATCTCTTGAGCATCTTTTGGCGACTGTCAAGTACTATGT 1980  
Db 1921 TTATGAGGCACTAAGGGGTATCTCTTGAGCATCTTTTGGCGACTGTCAAGTACTATGT 1980  
Qy 1981 CGAATCATCAAAACCAAGCTCCGCAATTCAGATTAATGTTTGTAAACGCTAGCGACCATGT 2040  
Db 1981 CGAATCATCAAAACCAAGCTCCGCAATTCAGATTAATGTTTGTAAACGCTAGCGACCATGT 2040  
Qy 2041 TCAAGAAACCAAAATGCTCAAGCTGATACCAATCAAAACGGAACCAAGCGAGAGAA 2100  
Db 2041 TCAAGAAACCAAAATGCTCAAGCTGATACCAATCAAAACGGAACCAAGCGAGAGAA 2100  
Qy 2101 ACCTCAGACAGAAACCTGAGGAAGAAACCCCTCGAGAGAGAAACCCGAAAGCGAGAA 2160  
Db 2101 ACCTCAGACAGAAACCTGAGGAAGAAACCCCTCGAGAGAGAAACCCGAAAGCGAGAA 2160  
Qy 2161 ACCAGAGTCTCAAAACCAACAGAGAGAAACCAAGAAATCACCAGAGGAATCAGAAGAAC 2220  
Db 2161 ACCAGAGTCTCAAAACCAACAGAGAGAAACCAAGAAATCACCAGAGGAATCAGAAGAAC 2220  
Qy 2221 TCAGTTCGAGCTGAAAAGTTGAAAGAAACCTGAGAGAGGCTGAAAGATTTACTTCGAAA 2280  
Db 2221 TCAGTTCGAGCTGAAAAGTTGAAAGAAACCTGAGAGAGGCTGAAAGATTTACTTCGAAA 2280  
Qy 2281 AATCCAGGAT 2290  
Db 2281 AATCCAGGAT 2290

RESULT 2  
US-09-536-784-65  
; Sequence 65, Application US/09536784  
; Patent No. 6573082  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/536,784  
; FILING DATE: 30-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: OCT-30-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Michelle S. Marks  
; REGISTRATION NUMBER: 41,971  
; REFERENCE/DOCKET NUMBER: PB340P3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512

```

; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2290 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-09-536-784-65

Query Match          99.8%; Score 2285; DB 4; Length 2290;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  TTGTTCCATGAACCTTGGTCGTACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGT 60
Db      1  TTGTTCCATGAACCTTGGTCGTACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGT 60

QY     61  TTCTTTATATAGATGGTGCATCAGGCTGTCTAAAAGCGCAGAAAACCTTGACACCAAGATGAAGT 120
Db     61  TTCTTTATATAGATGGTGCATCAGGCTGTCTAAAAGCGCAGAAAACCTTGACACCAAGATGAAGT 120

QY    121  CAGTAAAGAGGGAGGGGATCAACGCCGAACAAATNGTNATCAAGATTACGGATCAAGGTTA 180
Db    121  CAGTAAAGAGGGAGGGGATCAACGCCGAACAAATNGTNATCAAGATTACGGATCAAGGTTA 180

QY    181  TGTGACCTCTCATGAGAGACCATTAATCATTAATTAATGCGCAGAGTTCCTTATGATGCGCAT 240
Db    181  TGTGACCTCTCATGAGAGACCATTAATCATTAATTAATGCGCAGAGTTCCTTATGATGCGCAT 240

QY    241  CATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTATCAGTTGAAGGATTACAGACATTGT 300
Db    241  CATCAGTGAAGAGCTCCTCATGAAAGATCCGAATTATCAGTTGAAGGATTACAGACATTGT 300

QY    301  CAATGAAATCAAGGGTGGTTATGTCAATTAAGGTAAACGGTAAATATCTATGNTATACCTTAA 360
Db    301  CAATGAAATCAAGGGTGGTTATGTCAATTAAGGTAAACGGTAAATATCTATGNTATACCTTAA 360

QY    361  GGATCAGCTCATGCGGATTAATATTCCGACAAAGAAGAGATTAAACGTGAGAGCAGGA 420
Db    361  GGATCAGCTCATGCGGATTAATATTCCGACAAAGAAGAGATTAAACGTGAGAGCAGGA 420

QY    421  ACCCAGTCAATAATCATAACTCAAGAGCAGATATGCTGTTGCTGCAGCAGAGCCCCAAGG 480
Db    421  ACCCAGTCAATAATCATAACTCAAGAGCAGATATGCTGTTGCTGCAGCAGAGCCCCAAGG 480

QY    481  ACGTTATACAACGGATGATGGGTATATCTTCAATGCAATCTGATATCATTTGAGGACACGGG 540
Db    481  ACGTTATACAACGGATGATGGGTATATCTTCAATGCAATCTGATATCATTTGAGGACACGGG 540

QY    541  TGATGCTTATATCGTTCTCAGCGGACCATTAACCATTAATCTTAAAGGATCAGGTTATC 600
Db    541  TGATGCTTATATCGTTCTCAGCGGACCATTAACCATTAATCTTAAAGGATCAGGTTATC 600

QY    601  AGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAATGGGAAGCAGGGATCTCGTCCCTTC 660
Db    601  AGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAATGGGAAGCAGGGATCTCGTCCCTTC 660

QY    661  TTCAAGTTCTAGTTTATATGCAAAATCCAGCTCAACCAAGATTGTGAGAGAACCACAATCT 720
Db    661  TTCAAGTTCTAGTTTATATGCAAAATCCAGCTCAACCAAGATTGTGAGAGAACCACAATCT 720

QY    721  GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAACATTTCAAGCCCTTTACGTGA 780
Db    721  GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAACATTTCAAGCCCTTTACGTGA 780

QY    781  ATTGTATGCTTAAACCCCTTATCAGAAACGCCATGTGGAAATCTGATGGCCCTATTTTCGACCC 840
Db    781  ATTGTATGCTTAAACCCCTTATCAGAAACGCCATGTGGAAATCTGATGGCCCTATTTTCGACCC 840

QY    841  AGCGCAAAATCAAAAGTCGAACCGCGAGAGGTGTAGCTGTCCCTCATGGTAAACATTACCA 900
Db    841  AGCGCAAAATCAAAAGTCGAACCGCGAGAGGTGTAGCTGTCCCTCATGGTAAACATTACCA 900

```

1981	CGAACTCCAAACGAACGTCGCGATTTCAGATAATGTTTTGGTAACGCTAGCGCCATGT	2040
2041	TCAAAGAAACAAAAATGGTCAAGCTGATACCAATCAAACCGMAAAACCAAGCGAGAGAA	2100
2041	TCAAAGAAACAAAAATGGTCAAGCTGATCCATCAAACGMAAAACCAAGCGAGAGAA	2100
2101	ACCTCAGACAGAAAAACCTGAGGAGAGAAACCCCTCGAGAAAGAGAAACCGCAAAAGCGAGAA	2160
2101	ACCTCAGACAGAAAAACCTGAGGAGAGAAACCCCTCGAGAGAGAGAAACCGCAAAAGCGAGAA	2160
2161	ACCAGAGTCTCCAAAAACGAAACAGAGAGAACCGAAGAAATCACAGAGAGAAATCAGAGAAACC	2220
2161	ACCAGAGTCTCCAAAAACGAAACAGAGAGAACCGAAGAAATCACAGAGAGAAATCAGAGAAACC	2220
2221	TCAGGTTCGAGACTCAAAAGGTTTGAGAGAAAAACTCAGAGAGGCTCGAAGATTTACTTCGAAA	2280
2221	TCAGGTTCGAGACTCAAAAGGTTTGAGAGAAAAACTCAGAGAGGCTCGAAGATTTACTTCGAAA	2280
2281	AATCCAGGAT	2290
2281	AATCCAGGAT	2290

SULT 3  
-09-468-656A-11  
Sequence 11, Application US/09468656A  
Patent No. 6582706.  
GENERAL INFORMATION:  
APPLICANT: Johnson, Leslie S.  
APPLICANT: Adamou, John B.  
TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
TITLE OF INVENTION: Pneumoniae Group A and B Having Selected Structural  
TITLE OF INVENTION: Motifs  
FILE REFERENCE: 463201-444  
CURRENT APPLICATION NUMBER: US/09/468,656A  
CURRENT FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: 60/113,048  
PRIOR FILING DATE: 1998-12-21  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 2531  
TYPE: DNA  
ORGANISM: Streptococcus pneumoniae  
-09-468-656A-11

Query Match	97.0%;	Score	2220.2;	DB	4;	Length	2531;
Best Local Similarity	98.4%;	Pred. No.	0;				
Matches	2253;	Conservative	3;	Mismatches	16;	Indels	18;
						Gaps	1;
1	TTGTTCTTATGAAC	TTTGGTCTGTC	TACCAAGCTGGT	CAGGTTAAGAA	GAGTCTTAAT	TCAGT	60
57	TTGTTCTTATGAG	CTTGGAGCTT	ACCAAGCTGGT	CAGGATAAGAA	GAGTCTTAAT	TCAGT	116
61	TTCTTATATAGT	GTGATCAGGCTG	CTCAAGGCGAGAA	ACTTGGACCA	CAGATGAGT	120	
117	TGCTTATATAG	TNGTGTATCAGG	CTGTCAAAAGCGAGAA	ACTTGCAC	CCAGATGAGT	176	
121	CAGTAAAGAGG	AGGGATCAACG	CCGACAAATG	TGTTATCAAGAT	TACGGATCAAG	TTA	180
177	CAGTAAAGG	AGGGAGTCAACG	CCGACAAATG	TGTTATCAAGAT	TACGGATCAAG	TTA	236
181	TGTGACCTCT	CATGGAGAC	TATTAATCTATAAT	GCGAAGGTCCT	TATGATGCCAT	240	
237	TGTGACCTCT	CATGGAGAC	TATTAATCTATAAT	GCGAAGGTCCT	TATGATGCCAT	296	
241	CATCAGTGAAG	AGCTCCTCAT	TGAAGATCCGAAT	TATCAGTTGAAG	ATTTCAGACAT	TGT	300
297	CATCAGTGAAG	AGCTCCTCAT	TGAAGATCCGAAT	TATCAGTTGAAG	ATTTCAGACAT	TGT	356
301	CAATGAAAT	CAAGGGTGTTAT	GTCAATTAAGGTAA	ACGGTAAATAC	TATGTTACCTTAA	360	

357	DB	CAATGAAATCAAGGGTGGTTATGTCATTAAAGGTAAACGGTAAATACATATGTTTACCTTAA	411
361	QY	GGATGCAGCTCATCGGATATAATTCGGACAAAGAAGAGATTAAACGT CAGAAAGCAGGA	420
417	DB	:	
421	QY	GGATGCGCTCATCGGATATAATTCGGACAAAGAAGAGATTAAACGT CAGAAAGCAGGA	476
477	DB	ACGAGTCAATAATCAATACTCAAGAGCAGATAATGCTGTTGCTGCAGCCAGAGCCCAAGG	536
481	QY	ACGTTATACAAACGGATGATGGGTATATCTTCAATGCACTCTGATATCATTTGAGGACACGGG	540
537	DB	ACGTTATACAAACGGATGATGGGTATATCTTCAATGCACTCTGATATCATTTGAGGACACGGG	596
541	QY	TGATGCTTATATCGTTCTCAGCGGACCAATTACCAATTACATTCCTTAAGAAATGAGTTATC	600
597	DB	TGATGCTTATATCGTTCTCAGCGGACCAATTACCAATTACATTCCTTAAGAAATGAGTTATC	656
601	QY	AGCTAGCGATTAGTCTGCTCAGAAGCCTATTGGAAATGGGAAGCAGGATCTCGTCCCTTC	660
657	DB	AGCTAGCGAGTTAGTCTGCTCAGAAGCCTATTGGAAATGGGAAGCAGGATCTCGTCCCTTC	716
661	QY	TTCAAGTTCTAGTTATATGTCGAAATCCAGCTCAACCAAGATTGTCAGAGAACCAATCT	720
717	DB	TTCAAGTTCTAGTTATATGTCGAAATCCAGCTCAACCAAGATTGTCAGAGAACCAATCT	776
721	QY	GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACGTGA	780
777	DB	GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACGTGA	836
781	QY	ATTGTATGCTTAAACCCCTTATCAGAACGCCATGTGGAAATCTGATGGCCCTTATTTTCGACCC	840
837	DB	ATTGTATGCTTAAACCCCTTATCAGAACGCCATGTGGAAATCTGATGGCCCTTATTTTCGACCC	896
841	QY	AGCGAAAAATCAGAAATCGAACCAGAGAGTGTAGCTGTCCCTCATCGGTAACCAATTACCA	900
897	DB	AGCGAAAAATCAGAAATCGAACCAGAGAGTGTAGCTGTCCCTCATCGGTAACCAATTACCA	956
901	QY	CTTTATCCCTTTATGAAACAAATGTCGAAATTTGGAAAAACGAATTCCTCGTATTATTCCTCT	960
957	DB	CTTTATCCCTTTATGAAACAAATGTCGAAATTTGGAAAAACGAATTCCTCGTATTATTCCTCT	1016
961	QY	TCGTTATCGTTTCAAAACCAATTTGGGTACACAGATTCAGAGACCAAGAACCAAGTCCACAATC	1020
1017	DB	TCGTTATCGTTTCAAAACCAATTTGGGTACACAGATTCAGAGACCAAGAACCAAGTCCACAATC	1076
1021	QY	GACTCCGGAACTTAGTTCGAAGTCGGCAACTCTGCACCAAATCTCTCAACAGCTCCAGGAA	1080
1077	DB	GACTCCAGAACCTTAGTTCGAAGTCGGCAACTCTGCACCAAATCTCTCAACAGCTCCAGGAA	1118
1081	QY	TCCAAATTGATGAAATTTGGTTCAGAAAGCTGTTTCGAAAAAGTAGGCGATGCTTATGTCCTT	1140
1119	DB	TCCAAATTGATGGAATTTGGTTCAGAAAGCTGTTTCGAAAAAGTAGGCGATGCTTATGTCCTT	1178
1141	QY	TGAGGAGAAATGAGTTTCTCGTTTATATCCAGGCCAAGGATCTTTTCAGCAAGAAACAGCAGC	1200
1179	DB	TGAGGAGAAATGAGTTTCTCGTTTATATCCAGGCCAAGGATCTTTTCAGCAAGAAACAGCAGC	1238
1201	QY	AGGCATTTGATGAAACTTGCCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAA	1260
1239	DB	AGGCATTTGATGAAACTTGCCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAA	1298
1261	QY	AACGTACCTCCCATCTAGTGTATCGAAATTTTACAATAAGGCTTATGACTTACTAGCAAG	1320
1299	DB	AACGTACCTCCCATCTAGTGTATCGAAATTTTACAATAAGGCTTATGACTTACTAGCAAG	1358
1321	QY	AATTTACCAAGATTTACTTTGATATAAAGGTCCACAAGTTGATTTTTCAGGCGCTTTGGATAA	1380
1359	DB	AATTTACCAAGATTTACTTTGATATAAAGGTCCACAAGTTGATTTTTCAGGCGCTTTGGATAA	1418
1381	QY	CTGTTTGGAAACGACTCAAGGATCTCNCAAGTGATAAAGTCAAGTTAGTGGANGATATCTCT	1440
1419	DB	CTGTTTGGAAACGACTCAAGGATCTCTCAAGTGATAAAGTCAAGTTAGTGGANGATATCTCT	1478



897 CCAGCGCAATCACAAAGTCGAAACCGCAGAGGTGTAGCTGTCCCTCATGTAAACCAATTA 956  
898 CCACATTATCCCTTATGAAACAAATGCTGAATTTGAAAAACGAATTCGTGATTTATTC 957  
957 CCACATTATCCCTTATGAAACAAATGCTGAATTTGAAAAACGAATTCGTGATTTATTC 1016  
958 CTTTCGTTATCGTTTAAACCAATTTGGGTACCAAGTTCAGATTCAGAACCAACCAAGTTCACA 1017  
1017 CTTTCGTTATCGTTTAAACCAATTTGGGTACCAAGTTCAGATTCAGAACCAACCAAGTTCACA 1076  
1018 ATCGACTCCGGAACCTAGTCCAGTCCGCACTGCAACCAATCTCAACCAAGTTCACAG 1077  
1077 ATCGACTCCGGAACCTAGTCCAGTCCGCACTGCAACCAATCTCAACCAAGTTCACAG 1136  
1078 CAATCCAAATGATGAGAAATTTGCTCAAGAGAGCTTTGCAAAAGTAGCGCATGTTATGT 1137  
1137 CAATCCAAATGATGAGAAATTTGCTCAAGAGAGCTTTGCAAAAGTAGCGCATGTTATGT 1196  
1138 CTTTGAGGAAATGGAGTTTCTCGTTATATCCAGCAAGGATCTTTTACAGCAAAACAGC 1197  
1197 CTTTGAGGAAATGGAGTTTCTCGTTATATCCAGCAAGGATCTTTTACAGCAAAACAGC 1256  
1198 AGCAGGATTTAGTCAAACTGCGCAAGTCCGCAAGTTCATATCTCATAGCTAGAGCTAA 1257  
1257 AGCAGGATTTAGTCAAACTGCGCAAGTTCGCAAGTTCATATCTCATAGCTAGAGCTAA 1316  
1258 GAAAACTGACCTCCCATCTAGTGTAGTATGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 1317  
1317 GAAAACTGACCTCCCATCTAGTGTAGTATGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 1376  
1318 AAGAAATCCAAAGATTTACTTGAATAAAGTTCGCAAGTTCGATTTGAGGCTTTGGA 1377  
1377 AAGAAATCCAAAGATTTACTTGAATAAAGTTCGCAAGTTCGATTTGAGGCTTTGGA 1436  
1378 TAACTGTTGGAACGACTCAAGATGTCNCAAGTATAAAGTCAAGTATGAGGATAT 1437  
1437 TAACTGTTGGAACGACTCAAGATGTCNCAAGTATAAAGTCAAGTATGAGGATAT 1496  
1438 TCTTGCCTTTAGTCCGATTCGATTCATCAGAACGTTTGAAGAACCAATTCGCAAT 1497  
1497 TCTTGCCTTTAGTCCGATTCGATTCATCAGAACGTTTGAAGAACCAATTCGCAAT 1556  
1498 TACCTACTGATGATGAGATTCAGATTCAGATTCGCAAGTTCGCAAGTTCAGAACGATTC 1557  
1557 TACCTACTGATGATGAGATTCAGATTCAGATTCGCAAGTTCGCAAGTTCAGAACGATTC 1616  
1558 CGGTATATCTTTGATCCTGCTGATATACCAAGTATGAGGAGGATGCTTATGATTC 1617  
1617 CGGTATATCTTTGATCCTGCTGATATACCAAGTATGAGGAGGATGCTTATGATTC 1676  
1618 ACATATGACCCATAGCCATCGGATTAAGAAAGATGTTGCTGCAAGCTGAGAGAGCGC 1677  
1677 ACATATGACCCATAGCCATCGGATTAAGAAAGATGTTGCTGCAAGCTGAGAGAGCGC 1736  
1678 AGCCAGGCTTATGCTTAAAGAAAGGTTTGAACCTCTTCCACAGACCATCAGGATTC 1737  
1737 AGCCAGGCTTATGCTTAAAGAAAGGTTTGAACCTCTTCCACAGACCATCAGGATTC 1796  
1738 AGGAATATCTAGGCAAAAGGAGCAGAGCTATCTCAACCGGCTGAAAGCAGCTAAGAA 1797  
1797 AGGAATATCTAGGCAAAAGGAGCAGAGCTATCTCAACCGGCTGAAAGCAGCTAAGAA 1856  
1798 GGTGCACTGATGCTATGCTTCAATCTTCAATATCTGTAAGATCAAAACCGTAG 1857  
1857 GGTGCACTGATGCTATGCTTCAATCTTCAATATCTGTAAGATCAAAACCGTAG 1916  
1858 TTTAATCATACCTCATATGACCATACCATACCATCAATTTGATGTTGTTGCAAGAG 1917  
1917 TTTAATCATACCTCATATGACCATACCATACCATCAATTTGATGTTGTTGCAAGAG 1976  
1918 CTTTATGAGGCAACCTAAGGAGTATCTCTGAGGATCTTTTGGGACATGCTCAAGTACTA 1977

Db 1977 CTTTATGAGGCAACCTAAGGAGTATCTCTTGGGATCTTTTGGGACATGCTCAAGTACTA 2036  
QY 1978 TGTGAAACATCAAAACGAAAGTCCGATTCAGATTAATGTTTGGTAAACGCTAGCGACCA 2037  
Db 2037 TGTGAAACATCAAAACGAAAGTCCGATTCAGATTAATGTTTGGTAAACGCTAGCGACCA 2096  
QY 2038 TGTTCAAAGAAACAAAAATGTTCAAGCTGATACCAATCAAAACGAAAAACCAAGCGAGGA 2097  
Db 2097 TGTTCGTAATAAATAGGTAGACCAAGACAGTAAACCTGATGAAGATAAGGAACATGATGA 2156  
QY 2098 GAAACCTCAGACAGAAAAAACCTTGAGGAAGA-----AACCCC 2133  
Db 2157 AGTAAGTCAAGCAACTCACCTGTAATCTGTAAGAAAGAGAATCAAGCTGTTTAAATCC 2216  
QY 2134 TCGAGAGAGAAACCGCAAGGCGAGAAACAGAGTCTCCAAACCAACAGAGAGAACCGA 2193  
Db 2217 TTAGCAGATTAATCTTTTATAACCAAGCTGATACGGAAGAGACAGAGGAAGAGCTGA 2276  
QY 2194 AGAATCACCAGAGGATCAGAGAAACCTCAGCTCGAGACTGAAAGAGTTGAAGAAAAACT 2253  
Db 2277 AGATACCACAGATGAGGCTGAAATTCCTCAAGTAGAGAAATCTGTTATTAAACGCTAAGAT 2336  
QY 2254 GAGAGAGCTGAAGATTTACTTTGGAAAAATCCAGGAT 2290  
Db 2337 AGCAGATCGGAGGCTTGTAGAAAAAGTAACAGAT 2373

## RESULT 5

US-08-961-527-94  
; Sequence 94, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 94:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8195 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-527-94

Query Match 87.3%; Score 1999.2; DB 4; Length 8195;  
Best Local Similarity 92.1%; Pred. No. 0;  
Matches 2135; Conservative 1; Mismatches 154; Indels 27; Gaps 2;



QY 1 TTGTTCTTATGAACCTTGCTGTCACCAAGCTGGTCAGGTTAAAGAAAGCTCTAAATCGAGT 60  
Db 5658 TTGTTCTTATGAACCTTGCTGTCACCAAGCTGGTCAGGTTAAAGAAAGCTCTAAATCGAGT 5717  
QY 61 TTCTTTATAGATGGTGATCAGGCTGGTCAAAAGCGCAGAAAACCTTGACACCAAGATGAAGT 120  
Db 5718 TKCTTTATAGATGGTGATCAGGCTGGTCAAAAGCGCAGAAAACCTTGACACCAAGATGAAGT 5777  
QY 121 CAGTAAGAGGGAGGGATCAAGCGCGCAACAAATNGTNTATCAAGATTACGGATCAAGGTTA 180  
Db 5778 CAGTAAGAGGGAGGGATCAAGCGCGCAACAAATCGTCATCAAGATTACGGATCAAGGTTA 5837  
QY 181 TGTGACCTCTCATGAGACCAATATCATNTAATTAATGSCAAGGTTCTTATGATGCCAAT 240  
Db 5838 TGTGACCTCTCATGAGACCAATATCATNTAATTAATGSCAAGGTTCTTATGATGCCAAT 5897  
QY 241 CATCAGTCAAGAGCTCCATCAAGAGATCCGAATTTATCAGTTGAAGGATTCAGACATTTG 300  
Db 5898 CATCAGTCAAGAGCTCCATCAAGAGATCCGAATTTATCAGTTGAAGGATTCAGACATTTG 5957  
QY 301 CAATGAAATCAAGGTTGTTATGTCATTAAGGTAACCGTAAATATCTATGTTTACCTTAA 360  
Db 5958 CAATGAAATCAAGGTTGTTATGTCATTAAGGTAACCGTAAATATCTATGTTTACCTTAA 6017  
QY 361 GGATGCGAGCTCATGCGGATATATTCGGACAAAAGAGAGATTAAACGTCAGAAAGCAGGA 420  
Db 6018 GGATGCGAGCTCATGCGGATATATTCGGACAAAAGAGAGATTAAACGTCAGAAAGCAGGA 6077  
QY 421 ACGGAGTCATTAATCAACTCAAGAGC---AGATAATGCTGTTGCTGACGACAGGCCCA 477  
Db 6078 ACAAGTCATTAATCAAGGAGTGGTCTTAACGATCAAGCAGTAGTTGCGACGACAGGCCCA 6137  
QY 478 AGGACGTTATCAACGGATGATGGGTATATCTTCAATGCTGATCTGATATCATTTAGGACAC 537  
Db 6138 AGGACGTTATCAACGGATGATGGGTATATCTTCAATGCTGATCTGATATCATTTAGGACAC 6197  
QY 538 GGGTGATGCTTATATCGTTCTCAAGGACCAATACCAATTAATTCCTTAAAGATGAGTT 597  
Db 6198 GGGTGATGCTTATATCGTTCTCAAGGACCAATACCAATTAATTCCTTAAAGATGAGTT 6257  
QY 598 ATCAGTAGCGAGTTAGTCTGTCGAGAGCCATTGGAATGGGAAGCGAGGATCTCGTCC 657  
Db 6258 ATCAGTAGCGAGTTAGTCTGTCGAGAGCCATTGGAATGGGAAGCGAGGATCTCGTCC 6317  
QY 658 TTCTTCAAGTTCTAGTTATATGCAAAATCCAGCTCAACCAAGATTGTCAGAGAACCAAA 717  
Db 6318 TTCTTCAAGTTCTAGTTATATGCAAAATCCAGCTCAACCAAGATTGTCAGAGAACCAAA 6377  
QY 718 TCTGACTGTCTCAACTTATCATCAAAAATCAAGGGGAAAACATTTCAAGCCTTTTACG 777  
Db 6378 TCTGACTGTCTCAACTTATCATCAAAAATCAAGGGGAAAACATTTCAAGCCTTTTACG 7437  
QY 778 TGAATTTGATGCTAAACCTTATCAGAGCGCATGTGGAATCTGATGGCCTTATTTTCA 837  
Db 6438 TGAATTTGATGCTAAACCTTATCAGAGCGCATGTGGAATCTGATGGCCTTATTTTCA 6497  
QY 838 CCCAGCGCAATCAAGCTGCAAGCGCAGAGGCTGATGTCCTCTCATGTTAAACATTA 897  
Db 6498 CCCAGCGCAATCAAGCTGCAAGCGCAGAGGCTGATGTCCTCTCATGTTAAACATTA 6557  
QY 898 CCACCTTTATCCCTTATGAACAAATGCTGAAATGGAAAAACGAATGCTGCTATTATTCC 957  
Db 6558 CCACCTTTATCCCTTATGAACAAATGCTGAAATGGAAAAACGAATGCTGCTATTATTCC 6617  
QY 958 CTTTGGTTATCGTTCAAAACCAATGGGTACAGATTCAAGACAGAAACCAAGTCCACA 1017  
Db 6618 CTTTGGTTATCGTTCAAAACCAATGGGTACAGATTCAAGACAGAAACCAAGTCCACA 6677  
QY 1018 ATCGACTCCGNAACCTAGTCCAGTCCGCAACCTGCAACCAATCTCTCAACAGCTCCAA 1077  
Db 6678 ATCGACTCCGNAACCTAGTCCAGTCCGCAACCTGCAACCAATCTCTCAACAGCTCCAA 6737  
QY 1078 CAATCCAAATTGATGAGAAATTGGTCAAGAGAGCTGTTGGAAGTAGGCGATGTTATGT 1137

Db 6738 CAATCCAAATTGATGAGAAATTGGTCAAGAAAGCTGTTCCAAAAGTAGGCGATGTTATGT 6797  
QY 1138 CTTTGGAGGAGAAATGGAGTTTCTCGTTATATATCCAGCCAAAGGATCTTTTCAGCAGAAAACAGC 1197  
Db 6798 CTTTGGAGGAGAAATGGAGTTTCTCGTTATATATCCAGCCAAAGGATCTTTTCAGCAGAAAACAGC 6857  
QY 1198 AGCAGGCATTTGATAGCAAACTGGCCAAAGCAGGAGAAAGTTTATCTCATAGCTTAGGAGCTAA 1257  
Db 6858 AGCAGGCATTTGATAGCAAACTGGCCAAAGCAGGAGAAAGTTTATCTCATAGCTTAGGAGCTAA 6917  
QY 1258 GAAAACTGACCTCCCATCTAGTGATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 1317  
Db 6918 GAAAACTGACCTCCCATCTAGTGATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGC 6977  
QY 1318 AAGAAATTCACCAAGATTCTTCTGATATAAAGTCGACAAAGTTGATTTTGAAGGCTTTTGA 1377  
Db 6978 AAGAAATTCACCAAGATTCTTCTGATATAAAGTCGACAAAGTTGATTTTGAAGGCTTTTGA 7037  
QY 1378 TAACCTGTTGGAAACGACTCAAGGATGTCNCAAGTGTATAAAGTCAAGTTAGTGGANGATAT 1437  
Db 7038 TAACCTGTTGGAAACGACTCAAGGATGTCYCAAGTGATATAAGTCAAGTTAGTGGANGATAT 7097  
QY 1438 TCTTGGCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTATAGGAAAAACCAATGCGCAAT 1497  
Db 7098 TCTTGGCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTATAGGAAAAACCAATGCGCAAT 7157  
QY 1498 TACCTACACTGATGATGAGATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAGA 1557  
Db 7158 TACCTACACTGATGATGAGATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAGA 7217  
QY 1558 CGGTTATATCTTTGATCCTCGTGATATAACAGTGATGAGGGGATCGCTATGTAATCC 1617  
Db 7218 CGGTTATATCTTTGATCCTCGTGATATAACAGTGATGAGGGGATCGCTATGTAATCC 7277  
QY 1618 ACATATGACCCCATAGCCACTGGGATTTAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGC 1677  
Db 7278 ACATATGACCCCATAGCCACTGGGATTTAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGC 7337  
QY 1678 AGCCGAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTTCGACACACCAATAGGATTC 1737  
Db 7338 AGCCGAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTTCGACACACCAATAGGATTC 7397  
QY 1738 AGGAATACTGAGGCAAAAGGACAGACGATCTTCAACCGCGTGAAGCAGGCTAGAA 1797  
Db 7398 AGGAATACTGAGGCAAAAGGACAGACGATCTTCAACCGCGTGAAGCAGGCTAGAA 7457  
QY 1798 GGTGCCACTTTGATCGTATGCTTACCAATCTTCAATATATCTGTAGAAGTCAAAAAACGGTAG 1857  
Db 7458 GGTGCCACTTTGATCGTATGCTTACCAATCTTCAATATATCTGTAGAAGTCAAAAAACGGTAG 7517  
QY 1858 TTTAATCATACCTCATATATGACCAATTAACCAATACCAATTTGAGTGGTTTGAAGG 1917  
Db 7518 TTTAATCATACCTCATATATGACCAATTAACCAATTTGAGTGGTTTGAAGG 7577  
QY 1918 CTTTTATGAGGCACTTAAAGGGTATCTCTTGAGGATCTTTTGGCGACTCTCTCAAGTACTA 1977  
Db 7578 CTTTTATGAGGCACTTAAAGGGTATCTCTTGAGGATCTTTTGGCGACTCTCTCAAGTACTA 7637  
QY 1978 TGTGCAACATCCAAAACGACCGTCCGATTCAGATAATGGTTTTTGGTAAACGCTAGCAGCA 2037  
Db 7638 TGTGCAACATCCAAAACGACCGTCCGATTCAGATAATGGTTTTTGGTAAACGCTAGCAGCA 7697  
QY 2038 TGTTCAGAGAAACAAAATGGTCAAGCTGATACCAATCAAAACGGAAGAAACCAAGCAGGA 2097  
Db 7698 TGTTCAGAGAAACAAAATGGTCAAGCTGATACCAATCAAAACGGAAGAAACCAAGCAGGA 7757  
QY 2098 GAAACCTCAGACAGAAACCACTGAGGAAGA-----AACCCC 2133  
Db 7758 AGTAAGTGAAGCAACTCACCTGAAATCTGATGAAAAAGAGAATCACGCTGTTTAAATCC 7817  
QY 2134 TCGAGAGAGAGAAACCGCAAGCGAGAGAAACCAAGATGTCCTCAAAACCAACAGAGGAACAGA 2193

7818 TTCACGAGTAAATCTTTATTAATTAACCAAGCACTGATACGGAAGAGACAGAGGAGAAAGCTGA 7877

2194 AGAATCAACAGAGGAATCAGAGAACTCAGGTCGAGACTGAAAAGTTGAGAAAAAAT 2253

7878 AGATACCAAGATGAGGCTGAAATCCTCAAGTAGAGAAATCTGTTATTAAACGCTAAGAT 7937

2254 GAGAGAGGCTGAAGATTTTACTTCGAAAAAATCCAGGAT 2290

7938 AGCAGATGCGGAGGCCCTGCTAGAAAAAAGTAAACAGAT 7974

### RESULT 6

1-08-961-527-243  
 Sequence 243 Application US/08961527  
 Patent No. 6420135  
 GENERAL INFORMATION:  
 APPLICANT: Charles Kunach  
 TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 391  
 CORRESPONDENCE ADDRESS:

08-961-527-243

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,527  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 243:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2359 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

Query Match	60.1%;	Score 1376.6;	DB 4;	Length 2359;
Best Local Similarity	97.8%;	Pred. No. 0;		
Matches 1410;	Conservative 0;	Mismatches 14;	Indels 18;	Gaps 1;
1	TTGTTCCTATGAACTTCGTGCTCACCAGGTGGTCAGGTTAAGAAAGAGCTCAATCGAGT	60		
936	TTGTTCCTATGAGCTTGGAGCTTACCAAGCTGGTCAGGATPAGAAAGAGCTCAATCGAGT	995		
61	TTCTTATATAGATGGTGATCAGGTGGTCAAAAGCGCAGAAAACCTTGACACACAGATGAAGT	120		
996	TGCTTATATAGATGGTGATCAGGTGGTCAAAAGCGCAGAAAACCTTGACACACAGATGAAGT	1055		
121	CAGTAAAGAGGAGGGGATCAACGCCGACAAATNGTNNATCAAGATTACGGATCAAGGTGA	180		
1056	CAGTAAAGAGGAGGGGATCAACGCCGACAAATNGTNNATCAAGATTACGGATCAAGGTGA	1115		
181	TGTGACCTCTCATGGAGACCATATCATCTATTAATGGCAAGGTCCTTATGATGCCAT	240		
1116	TGTGACCTCTCATGGAGACCATATCATCTATTAATGGCAAGGTCCTTATGATGCCAT	1175		

QY	241	CATCAGTGAAGAGCTCCTCATGAAGAAGATCGGAATTTATCAGTTGCAAGATTTCAGACATTGT	300
Db	1176	CATCAGTGAAGAGCTCCTCATGAAGAAGATCGGAATTTATCAGTTGCAAGATTTCAGACATTGT	1235
QY	301	CAATGAAATCAAGGGTGGTTATGTCATTAAAGTAAACGGTAATAACTATATGENTACCTTAA	360
Db	1236	CAATGAAATCAAGGGTGGTTATGTCATTAAAGTAAACGGTAATAACTATATGTTACCTTAA	1295
QY	361	GGATGCAGCTCATGCGGATAATATTCGGAACAAAGAGAGATTAAACGTCAGAAAGCAGGA	420
Db	1296	GGATGCAGCTCATGCGGATAATATTCGGAACAAAGAGAGATTAAACGTCAGAAAGCAGGA	1355
QY	421	ACGAGTCAATAATCAATCACTCAAGACAGATATGCTGTGGTCGACGCCAGAGCCCAAGG	480
Db	1356	ACGAGTCAATAATCAATACTCAAGAGCAGATAATGCTGTGTCGACGCCAGAGCCCAAGG	1415
QY	481	ACGTTATACAACGGATGATGGGTATATCTTTCAATGATCTGATATCATTTACAGAGACAGGG	540
Db	1416	ACGTTATACAACGGATGATGGGTATATCTTTCAATGATCTGATATCATTTACAGAGACAGGG	1475
QY	541	TGATGCTTATATCGTTCTCTCAGCGGACCAATTACCAITTAATTCTTAAGAAATGAGTTATC	600
Db	1476	TGATGCTTATATCGTTCTCTCAGCGGACCAATTACCAITTAATTCTTAAGAAATGAGTTATC	1535
QY	601	AGCTAGGGATTAGCTGTCGAGAAGCTTATGGAAATGGGAAGCAGGGATCTCGTCTCTTC	660
Db	1536	AGCTAGCGAGTTAGCTGCTGAGAAGCTTATGGAAATGGGAAGCAGGGATCTCGTCTCTTC	1595
QY	661	TTCAAGTCTCAGTTATAATGCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACGTGA	720
Db	1596	TTCAAGTCTCAGTTATAATGCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACGTGA	1655
QY	721	GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACGTGA	780
Db	1656	GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAATTTCAAGCCCTTTTACGTGA	1715
QY	781	ATTGTATGTAAACCTTATCAGAACGCCATGTGGAATCTGATGGGCTTATTTTCGACCC	840
Db	1716	ATTGTATGTAAACCTTATCAGAACGCCATGTGGAATCTGATGGGCTTATTTTCGACCC	1775
QY	841	AGCGCAATTCAGAAGTCGAACGCCAGAGGTGTAGCTGTCCCTCATGTTAAACCATTACCA	900
Db	1776	AGCGCAATTCAGAAGTCGAACGCCAGAGGTGTAGCTGTCCCTCATGTTAAACCATTACCA	1835
QY	901	CTTTATCCCTTATGAACAAATGTCTGAAATGGAAAAACGAAATTCGCTGATTTATTCGCCCT	960
Db	1836	CTTTATCCCTTATGAACAAATGTCTGAAATGGAAAAACGAAATTCGCTGATTTATTCGCCCT	1895
QY	961	TCGTTATCGTTCAAAACATTTGGGTACAGATTCAAGACCAAGAACCAAGTCCCAACAC	1020
Db	1896	TCGTTATCGTTCAAAACATTTGGGTACAGATTCAAGACCAAGAACCAAGTCCCAACAC	1955
QY	1021	GACTCGGAACCTTAGTCCCAAGTCGCAACCTGCACCAATCTCTCAACCGAGCTCCAGGAA	1080
Db	1956	GACTCCAGAACCTTAGTCCCAAGTCG-----CAACCGAGCTCCAGGAA	1997
QY	1081	TCCAAATGATGAAATTCGTCAAAGAGCTGTTTCGAAAAGTAGGCGATGGTTATCTCTT	1140
Db	1998	TCCAAATGATGAAATTCGTCAAAGAGCTGTTTCGAAAAGTAGGCGATGGTTATCTCTT	2057
QY	1141	TGAGGAAATGGAGTTTCTGTTTATATCCAGCCAGGATCTTTTCAGCAGAAACAGCAGC	1200
Db	2058	TGAGGAAATGGAGTTTCTGTTTATATCCAGCCAGGATCTTTTCAGCAGAAACAGCAGC	2117
QY	1201	AGGCATTGTAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTTAAGAA	1260
Db	2118	AGGCATTGTAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTTAAGAA	2177
QY	1261	AACTGACCTCCCATCTAGTGATCGAGAAATTTTACAAATAGGCTTATGACTACTACGAAG	1320
Db	2178	AACTGACCTCCCATCTAGTGATCGAGAAATTTTACAAATAGGCTTATGACTACTACGAAG	2237
QY	1321	AAPTCACCAAGATTTACTTTGATAATAAAGGTCGAACAAAGTTGATTTTGAAGGCTTTGGATAA	1380



1516 GATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAAAGACGGTTATATCTTTTGATCC 1575  
1602 AGTTGCTGTTGTTCAATAGCTGATAGTATACACAGCTCAGATGGTTATATTTGATGA 1661  
1576 TCGTGATATACAGATGATGAGGGGATGCTATGTAACTCCACATAGACCCATAGCCA 1635  
1662 ACATGATATATATCAGTGAATGAGGAGATGATATGTAACGCTCATATGGGCCATAGTCA 1721  
1636 CTGGATTATATATATATGTTGCTGAGCTGAGAGAGCGGACCCAGGCTTATGCTAA 1695  
1722 CTGATTGGAAGATAGACCTTTCTGATAGGAAAGATTGACGCTCAAGGCTATATCTAA 1781  
1696 AGAGAAAGGTTTGACCCCTCTCTCCAGACACCATCAGGATTACAGAAATATCTGAGGCAA 1755  
1782 AGAAAGGTTATCTTACTCTTCTCCAGAGCGAGATGTTAAAGCAATCAACTGGGA 1841  
1756 AGAGCAAGAGCTATCTACACCGCGTGAAGACGCTAAGAAAGTGCCACTTGTATCGTAT 1815  
1842 TAGTGACGAGCTATTTACAAATCGTGTGAAGGGGAAAAACGAATTCACCTCGTTGACT 1901  
1816 GCCTTACATCTTCAATATATCTAGAGTCAAAACGGTATGTTAATCATACCTCATTA 1875  
1902 TCCATATATGTTGAGCATACAGTTGAGGTTAAAGCGGTAATTTGATTTCTCTCTAA 1961  
1876 TGACCATATACCAATCAATCAATTTGAGTGGTTTGAAGAGGCGCTTTATAGGCACTTAA 1935  
1962 GGATCATACCAATATATTAATTTGCTTGGTTGATGATCACACATCAAGCTCCAA 2021  
1936 GGGGTATATCTTTGAGGATCTTTTGGCGACTGTCAGTACTATCTCGAATCACTCAACGA 1995  
2022 TGGCTATATCTTGAAGATTTGTTGCGAGATTAGTACTACGTAGAACACCTTGACGA 2081  
1996 ACCTCCGATTCAGATATGTTTGTAGCTGAGGACCATGTTTCAAGAAACAAAA 2055  
2082 ACCTCCACATCTTAATGATGATGGGCAATGCGCAGTGAGCATGTGTTAGGCAAGAA 2141  
2056 TGCTCAAGCTGATACCAATCAACGGAAGAAACCAAGCGAGGAGAACTTCAGACAGAAA 2115  
2142 CCACAGTGAAGTCCAAATTAAGAACTTCAAGCGGATGAAGGCGCATAGAGGAAACACC 2201  
2116 ACCTGAGGAAGAAACCCCTCGAGAGAGA 2144  
2202 TGCTGAGCCAGAGTCCCTCAAGTAGAGA 2230

## RESULT 8

US-08-961-083-55

Sequence 55, Application US/08961083

Patent No. 6159469

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,083

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-083-55

Query Match 43.1%; Score 987.6; DB 3; Length 2389;  
Best Local Similarity 67.7%; Pred. No. 8.3e-293;  
Matches 1481; Conservative 0; Mismatches 645; Indels 60; Gaps 5;  
QY 4 TTCCTATGAACCTTGTCGTCACCAAGCTCGTCAAGAGGTTAAGAAAGAGTCTAATCGAGTTTC 63  
Db 1 TTCCTACGATTGGGACTGTATCAAGCTAGAACGGTTAAGGAAAA---TATCTGTTTC 57  
QY 64 TTATATAGATGTCGTCAGGCTCGTCAAAAGGCAGAAAACTTCACACACAGATGAAGTCAG 123  
Db 58 CTATATAGATGAARAACAGCGCAGCAAAACCGGAGATTGACTCTCTGATGAGGTAG 117  
QY 124 TAAGAGGAGGGATCAACGCCGAAACAATNGTNATCAAGATTACGGATCAAGGTTATGT 183  
Db 118 CAAGCGTGAAGGAATCAATGCTGAGCAAAATCGTCATCAAGATAACAGACCAAGGCTATGT 177  
QY 184 GACCTCTCATGAGACCATTATCAATCTATTAATGCGAAGGTTCTTATGATGCCATCAT 243  
Db 178 CACTTCATGCGGACCACTCATTTATTAATGTAAGGTTCTTATGACGCTATCAT 237  
QY 244 CAGTCAAGAGCTCCTCATGAAGATCCGAATTTATCAGTTCAAGGATTACACATTTGTCAA 303  
Db 238 CAGTGAAGATTTACTCATGAAGATCCAACTATAAGCTAAGATGAGATATTGTAA 297  
QY 304 TGAATCAAGGTGTTATGTCTATTAAGGTAAACGGTAAATCTATGTTNACCTTAAGGA 363  
Db 298 TGAGGTCAAGGTGATATGTTTATCAAGGTAGTGAATACTATGTTTACCTTAAGGA 357  
QY 364 TGCAGCTCATCGGATATATTCGACAAAGAGAGATTAACTGTCAGAGAGAGAACG 423  
Db 358 TGCTGCCACCGGATAACGTCCTGACAAAAGAGGAAATCAATCGACAAAAACAAGAGCA 417  
QY 424 CAGTCAATATCAT-----AACTCAAGAGCAGATAATGCTGTTGTCGACCGACAGC 474  
Db 418 TAGTCAACATCGTGAAGGTGGAACTCCAAGAACGATGCTGTTGCTTGGCAGCTTC 477  
QY 475 CCAAGGAGTTATACACGGATGATGGGTATATCTTCAATGATCTGATATCAATGAGGA 534  
Db 478 GCAAGGAGCTATATACTACAGATGATGTTATATCTTTAATGCTTCTGATATCATAGAGA 537  
QY 535 CACGGTGATGCTTATATCTTCCTCAACGCAACCATTAACCATTAATCTTCTTAAGATGA 594  
Db 538 TACTGTGATGCTTATATCGTTCTCATGGAGATCATTAACCATTAATCTTCTTAAGATGA 597  
QY 595 GTTATCAGCTAGCGATTAGCTGCTGCAAGAGCCT----- 629  
Db 598 GTTATCAGCTAGCGATTGGCTGCTGCAAGAGCCTTCTTATCTGGTCGAGGAATCTGTC 657  
QY 630 -----ATTGGAATGGGAAGCAGGATCTCGTCTTCTTCAAGTTCTAGTTTATATGC 681  
Db 658 AAATTCAGAAACCTATCGCGACAAAATAGCGATAACACTTCAAGAAACAACTGGGTACC 717  
QY 682 AAATCCAGCTCAACCAAGATTGTGAGAGAACCAATCTGACTGTCTCACTTCACTTATCA 741  
Db 718 TTCTGTAGCAATCCAGGAACCTACAAATACTAACCAAGCAACACAGCAACACTAACAG 777  
QY 742 TCA---AAATCAAGGGGAAACATTTCAAGCCTTTTCAAGTATGTTATGCTTAAACCCCT 798

778 TCAAGCAAGTCAAAGTAAATGATAGTCTCTTGAACAGCTCTACAACTGCCTTT 837  
799 ATCAGAAACGCAATGGAATCTGATGGCTTATTTTCAGCCAGCGCAAAATCACAAAGTCG 858  
838 GAGTCAACGACATGATAGTCTGATGGCTTGTCTTTGATCCAGCAAAATCACAAAGTCG 897  
859 AACCGCAGAGGTGATGCTGCCCTCATGGTAAACATTAACACATTTATCCCTTATGAACA 918  
898 AACAGCTAGAGGTGTGAGTGGCCACACGCGAGATCATTAACCATTTATCCCTTATCTCCA 957  
919 AATGCTGAATTTGAAAAACGAATTTGCTGATATATTCCTTATTCCTTATCGTTTCAAAACA 978  
958 AATGCTGAATTTGAGAACGATCTGCTGATATATTCCTTATTCCTTATCGTTTCAAAACA 1017  
979 TTGGGTACCAAGATTCAAGACAGAACCAACCAAGTCCCAATCGATCCGGAACCTAGTCC 1038  
1018 TTGGGTACCAAGATTCAAGCGCAGAACCAACCAAGTCCCAACCGACTCCGGAACCTAGTCC 1077  
1039 AAGTCCGCAACCTGCACCAAAATCTCAACGAGCTCCAAGCAATCCAATTTGATGAAATTT 1098  
1078 AGGCCCGCAACCTGCACCAAAATCTTAAATAGACTCAA-----ATTCTCTTTT 1125  
1099 GGTCAAAGAGCTGTTCGAAAGTAGGCGATGGTTATGTTTGGAGGAATGGAGTTTC 1158  
1126 GGTATGCTGATGCTGACGAAAGTTGGGAAGGATATGTTTTCGAAAGAGGGGATCTC 1185  
1159 TCGTTATATCCAGCCAGGATCTTTTCAGCGAAACAGCAGCAGGCAATTTAGTACAACT 1218  
1186 TCGTTATGCTTTTCGAAAGATTACCATCTGAAACTGTAAAATCTTTGAAAGCAAGTT 1245  
1219 GGCCAGCAGGAAAGTTTATCTCATAGCTAGGCTAAGAACTGACCTCCCATCTAG 1278  
1246 ATCAAAACAGAGAGTGTTCACACACTTTAAGTCTTAAAGAAAGAAATGTTGCTCTCTG 1305  
1279 TGATCGAGAAATTTACATAAGGCTTATGACTTACTAGCAAGAAATTCACCAAGATTTACT 1338  
1306 TGACCAAGAAATTTATGATAAGCAATATACTGTTAACTGAGGCTCAAGAGCTTGT 1365  
1339 TGATAATAAGGCTCGCAAGTTGATTTTTCAGGCTTTGGATTAACCTGTTGGAACGACTCAA 1398  
1366 TGAATAAGGCTGTAATCTGATTTCCAGGCTTAGCAAAATTTATGAAAGCTTGA 1425  
1399 GGATGTCNCAAGTGAATGAAGTCAAGTTAGTGGAGATATCTTGCTCTTTAGTCCGAT 1458  
1426 TGATGAATCGACTAATAAGAAAGAAATTTGATGATGATTTATTTGGCAATCTAGCAACCAAT 1485  
1459 TCGTATCCCAAGCTTTAGGAABACCAATTCGCAAAATTTACCTACATGATGATGAGAT 1518  
1486 TACCATCCAGAGGCTTTGGCAACCAAAATTTCTCAAAATTTGATGATGATGATGAGAGT 1545  
1519 TCAAGTAGCCAAAGTTGGCAGCAAGTACACAAAGAGACGGTTATATCTTTGATCTCTG 1578  
1546 TCGTATTTGCTCAATTTGATGATAGTATACACGCTCAGATGTTTATATTTTATGTAACA 1605  
1579 TGATATAACAGTATGAGGGGATGCTTATGTAATCCACATATGACCAATAGCCACTG 1638  
1606 TGATATAATCAGTATGAGGAGATGATATGTAACGCTCATATGGGCCATAGTCACTG 1665  
1639 GATTAAAAAGATAGTTTGTCTGAGCTGAGAGCGGCGAGCCAGGCTTATGCTTAAGA 1698  
1666 GATTGGAAGATAGCTTTCTGATAAGGAAAGTTTCAGCTCAAGCTTATCTACTATAAGA 1725  
1699 GAAAGTTTGAACCTCTCTGACAGACATCAGGATTCAGGAATATCTGAGGCAAAAGG 1758  
1726 AAAAGGATCTTACTCTCAATCTCCAGAGCGAGATGTTTAAAGCAATTCCTACTGGAGATAG 1785  
1759 AGCAGAGCTATCTACAAACCGGCTGAAAGCAGCTAAGAGGTCGCCATTTGATCGTATGCC 1818  
1786 TGCAGCAGCTATTTACATCTGTTGAAAGGGGAAACGAATTTCCACTCGTTCGACTTCC 1845  
1819 TTCAATCTCAATATATCTAGTAAGAGTCAAAACCGGTAGTTTAAATCACTACATTAAGA 1878  
1846 ATATATGTTGAGCATACAGTTGAGTTTAAAGAAACGTTAATTTGATTTATTTCTCTATAAGA 1905

QY 1879 CCATTACCATTAATCAATATTTGAGTGGTTTGAAGAGCCTTTATGAGGACCTTAAGGG 1938  
Db 1906 TCATTACCATTAATATTAATTTGTTGGTTTGTATGATCACACATACAAAGCTCCAAATGG 1965  
QY 1939 GTATATCTTTGAGGATCTTTTGGCGACTGTCAAGTACTATGTGGAACATCCAAACGAACG 1998  
Db 1966 CTATACCTTTGGAAGATTTGTTTGCAGCATTAAGTACTAGTAGAACACCTTGACGAACG 2025  
QY 1999 TCCGATTCAGATATATGTTTGGTAAACGTACGACCATGTTTCAAGAAACAAAAATGG 2058  
Db 2026 TCACATTTCTAATGATGGATGGGCAATGCCATGAGCATGTGTAGGCAAGAAGACCA 2085  
QY 2059 TCAAGCTGATACCAATCAACGGAAGAAACCAACGCGGAGAAACCTTCAGACAGAAAAACC 2118  
Db 2086 CAGTGAAGATCAAAATTAAGAACTTCAAGCGGATGAAGAGCCAGTAGAGGAAACACCTGC 2145  
QY 2119 TGAGGAGAAACCCCTCGAGAGAGA 2144  
Db 2146 TGAGCGAAGATCCCTCAAGTAGAGA 2171

## RESULT 9

US-09-536-784-55

; Sequence 55, Application US/09536784

; Patent No. 6573082

; GENERAL INFORMATION:

; APPLICANT: Choi et. al.

; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines

; NUMBER OF SEQUENCES: 452

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/536,784

; FILING DATE: 30-Oct-1997

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/961,083

; FILING DATE: OCT-30-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Michelle S. Marks

; REGISTRATION NUMBER: 41,971

; REFERENCE/DOCKET NUMBER: PB340P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 55:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2389 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 55:

US-09-536-784-55

## Query Match

43.1%; Score 987.6; DB 4; Length 2389;

Best Local Similarity 67.7%; Pred. No. 8.3e-293;

Matches 1481; Conservative 0; Mismatches 645; Indels 60; Gaps 5;

QY 4 TTCTATGACTTGGTTCGTCACCAAGCTGTCAGGTTAAGAAAGTCTTAATCGAGTTTC 63

Db 1 TTCTTACGAGTGGGACTGTATCAAGCTAGAACGGTTAAGGAAAA---TAATCGTGTTC 57



US-08-961-527-355  
; Sequence 355, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 355:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 973 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-527-355

Query Match 34.7%; Score 794.8; DB 4; Length 973;  
Best Local Similarity 98.7%; Pred. No. 8.1e-234;  
Matches 810; Conservative 0; Mismatches 10; Indels 1; Gaps 1;  
QY 1 TGTTCCTATGAATGGTGGTCACCAAGCTGTGAGGTTAAGAAAGAGTCTAATCGAGT 60  
DB 154 TGTTCCTATGAGCTTGGACGTTACCAAGCTGTGAGGTTAAGAAAGAGTCTAATCGAGT 213  
QY 61 TCTTTATATAGATGGTGATCAGGCTGGTCAAAAGGCGAGAAACTTGACACCCAGATGAAGT 120  
DB 214 TGTTCCTATATAGATGGTGATCAGGCTGGTCAAAAGGCGAGAAACTTGACACCCAGATGAAGT 273  
QY 121 CAGTAAGAGGGAGGGGATCAAGCCGAAACAATNGTNATCAAGATTACGGATCAAGGTTA 180  
DB 274 CAGTAAGAGGGAGGGGATCAAGCCGAAACAATNGTNATCAAGATTACGGATCAAGGTTA 333  
QY 181 TGTGACCTCTCATGAGACCATTCATCTATCTAATGTCGAGGTTCCCTATGATGCCAT 240  
DB 334 TGTGACCTCTCATGAGACCATTCATCTAATGTCGAGGTTCCCTATGATGCCAT 393  
QY 241 CATCATGTAAGAGCTCCTCATGAAGATCCGAAATTCATGTTCAAGGATTCAGACATTTGT 300  
DB 394 CATCATGTAAGAGCTCCTCATGAAGATCCGAAATTCATGTTCAAGGATTCAGACATTTGT 453  
QY 301 CAATGAATCAAGGGTGGTTATGTCTAATGAAGTAAACGGTAAATACTATGTTNACCTTAA 360  
DB 454 CAATGAATCAAGGGTGGTTATGTCTAATGAAGTAAACGGTAAATACTATGTTNACCTTAA 513  
QY 361 GGATGAGCTCATGGGATTAATATTCGGGCAAAAGAGAGATTAAACGGTCAGAGCAGGA 420  
DB 514 GGATGAGCTCATGGGATTAATATTCGGGCAAAAGAGAGATTAAACGGTCAGAGCAGGA 573

QY 421 AGCAGTCTATATCATTAACCTCAAGAGCAGATATGCTGTTGCTGCAGCCAGAGCCCAAGG 480  
DB 574 AGCAGTCTATATCATTAACCTCAAGAGCAGATATGCTGTTGCTGCAGCCAGAGCCCAAGG 633  
QY 481 ACCTTATACAAACCGATGATGGGTATATCTTCAATGATCATCTGATATCATTTGAGGACACGGG 540  
DB 634 ACCTTATACAAACCGATGATGGGTATATCTTCAATGATCATCTGATATCATTTGAGGACACGGG 693  
QY 541 TGATGCTTATATCGTTCTCTCACGGGACCAATACCAATTACCTTCAAGAAATGAGTTATC 600  
DB 694 TGATGCTTATATCGTTCTCTCACGGGACCAATACCAATTACCTTCAAGAAATGAGTTATC 753  
QY 601 AGCTAGCGAGTTAGCTGCTGCAGAGGCTATTGGAATGGGAAGCAGGATCTGCTCCTTC 660  
DB 754 AGCTAGCGAGTTAGCTGCTGCAGAGGCTATTGGAATGGGAAGCAGGATCTGCTCCTTC 813  
QY 661 TTCAAGTTCTAGTTTATAATGCAATCCAGTCAACCAAGATTGTGAGAAACCAACAATCT 720  
DB 814 TTCAAGTTCTAGTTTATAATGCAATCCAGTCAACCAAGATTGTGAGAAACCAACAATCT 873  
QY 721 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAACATTTCAAGCCTTTTACGTGA 780  
DB 874 GACTGTCACTCCAACTTATCATCAAAATCAAGGGGAAAAACATTTCAAGCCTTTTACGTGA 933  
QY 781 ATTGTATGCTAAACCTTATCATCAAGCCGATGTGGAATCTG 821  
DB 934 ATTGTATGCT-AAACCTTATCATCAAGCCGATGTGGAATCTG 973

## RESULT 11

US-08-961-527-258  
; Sequence 258, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 258:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1684 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-527-258  
Query Match 29.6%; Score 677; DB 4; Length 1684;  
Best Local Similarity 99.9%; Pred. No. 1.9e-197;



Matches 688; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1602 ATCCCTATGTAACCTCCACATATGACCATAGCCACTGGATTAAAGAGATGTTGTCTG 1661  
Db 1 ATGCTTATGTAACCTCCACATATGACCATAGCCACTGGATTAAAGAGATGTTGTCTG 60  
QY 1662 AAGCTGAGAGAGGGGAGCCGAGCTTATGCTAAAGAGAGTTTGAACCCCTCCTTCGA 1721  
Db 61 AAGCTGAGAGAGGGGAGCCGAGCTTATGCTAAAGAGAGTTTGAACCCCTCCTTCGA 119  
QY 1722 CAGACCATCAGGATTCAGGAAATCTAGGCAAAATCTAGGCAAAAGGAGCAGTCTTACAAACGGG 1781  
Db 120 CAGACCATCAGGATTCAGGAAATCTAGGCAAAAGGAGCAGTCTTACAAACGGG 179  
QY 1782 TGAAGCAGCTAAGAGGTCCTTATGCTATGCTTCAATCTTCAATATATCTGTAG 1841  
Db 180 TGAAGCAGCTAAGAGGTCCTTATGCTATGCTTCAATCTTCAATATATCTGTAG 239  
QY 1842 AAGTCAAAACGGTAGTTAATCATACCTCATTTATGACCAATTCATTAACCAATTTG 1901  
Db 240 AAGTCAAAACGGTAGTTAATCATACCTCATTTATGACCAATTCATTAACCAATTTG 299  
QY 1902 AGTGGTTTGAAGAGGCTTATGAGGACCTTAAGGGTATATCTTTGAGGATCTTTGG 1961  
Db 300 AGTGGTTTGAAGAGGCTTATGAGGACCTTAAGGGTATATCTTTGAGGATCTTTGG 359  
QY 1962 CGACTCTCAAGTACTATGTGCAATCCAAACGCTCCGCACTTCAAGTAAATGTTTGG 2021  
Db 360 CGACTCTCAAGTACTATGTGCAATCCAAACGCTCCGCACTTCAAGTAAATGTTTGG 419  
QY 2022 GTAACCTGAGGACCTTGTCAAGAGAAACAAATGCTCAAGTGTATCCAAATCAACGG 2081  
Db 420 GTAACCTGAGGACCTTGTCAAGAGAAACAAATGCTCAAGTGTATCCAAATCAACGG 479  
QY 2082 AAAACCAAGCAGGAGAACTTCAAGAGAAACAAATGCTCAAGTGTATCCAAATCAACGG 2141  
Db 480 AAAACCAAGCAGGAGAACTTCAAGAGAAACAAATGCTCAAGTGTATCCAAATCAACGG 539  
QY 2142 AGAAGCCGAAAGCAGGAGAACTTCAAGAGAAACAAATGCTCAAGTGTATCCAAATCAACGG 2201  
Db 540 AGAAGCCGAAAGCAGGAGAACTTCAAGAGAAACAAATGCTCAAGTGTATCCAAATCAACGG 599  
QY 2202 CAGAGGAATCAGAGAACTTCAAGAGAACTTCAAGAGAACTTCAAGAGAACTTCAAGAGG 2261  
Db 600 CAGAGGAATCAGAGAACTTCAAGAGAACTTCAAGAGAACTTCAAGAGAACTTCAAGAGG 659  
QY 2262 CTGAAGATTCTTGAAGAAATCCAGAT 2290  
Db 660 CTGAAGATTCTTGAAGAAATCCAGAT 688

## RESULT 12

US-08-961-083-181  
; Sequence 181, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,083

; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 181:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-961-083-181

Query Match 17.5%; Score 401; DB 3; Length 1342;  
Best Local Similarity 65.7%; Pred. No. 1.1e-112;  
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;  
QY 23 CACCAAGCTGCTCAGGTTAAAGAAAGAGTCTAATCGAGTTTCTTATATAGATGTTGATCAG 82  
Db 8 CAGCATCGTTCGACGAGAAATAAGGACAATAATCGTGTCTCTTATGTGGATGGCAGCCAG 67  
QY 83 GCTGTTCAAAAGGCGAGAAACCTTGACACAGATGAAGTCAGTAAGAGGGAGGGGATCAAC 142  
Db 68 TCAAGTTCAGAAAGTGAAGAACTTGACCCAGACCCAGGTTAGCCAGAAAGAAAGAAATTCAG 127  
QY 143 GCCGAACAAATGTTNATCAAGATTACGGATTCAAGGTTATGTGACCTCTCATGGAGACCAT 202  
Db 128 GCTGAGCAAAATGTTNATCAAAATTCAGATCAGGCTATGTAACTGTCACCGGTGACCCAC 187  
QY 203 TATCATTTACTATATGCGCAAGGTTCCCTTATGATGCCATCATCAGTGAAGAGTCCCTCATG 262  
Db 188 TATCATTTACTATATGCGCAAGGTTCCCTTATGATGCCCTCTTATGTGAAGAACTCTTGATG 247  
QY 263 AAAGATCCGAATTTACGTTGAAGGATTCAGACATTTGTCATCAATGAAATCAAGGGTGGTTAT 322  
Db 248 AAGGATCCAAACTATCAACTTAAAGACGCTGATATTTGTCATGAAGTCAAGGGTGGTTAT 307  
QY 323 GTCATTAAAGTAAAGGTTAAATCTATGTTNATACCTTAAAGATGTCAGCTCATCGGATAT 382  
Db 308 ATCATCAAGGTCGATGGAATAATTTATGTTCTACCTGAAAGATGTCAGCTCATGCTGATAAT 367  
QY 383 ATTCGGACAAAGAGAGATTAAAGCTCAGAGCAGGAAACGAGTCATAATCATAACTCA 442  
Db 368 GTTCGAACCTAAAGATGAATCAATGTCATAAACAAGAAC---ATGTCAAGATATATGAG 424  
QY 443 AGAGCAGATAATGCTGTTGCTGACGCCAGAGCCCAAGGACGTTTATACACGGATATGAGG 502  
Db 425 AAGGTTAACTCTAATGTTGCTGTAGCAAGGCTCTCAGGGACGATATACGACAAATGATGTT 484  
QY 503 TATATCTTCAATGATCTGATATCAATTTGAGGACAGCGGTTGATCTTATATCGTCTCTCAC 562  
Db 485 TATGTTTAAATCCAGCTGATATTTATCGAAGATACGGGTAATGCTTATATCGTCTCTCAT 544  
QY 563 GCGACACCATTTACCATTTACATTTCCCTAAGATGATGTTATCAGCTAGCGAGTTAGCTGTGCA 622  
Db 545 GGAGGTCACTATCACTATCTATCTCCCAAAAGCGATTTATCTGCTAGTGAATTAGCAGAGCT 604  
QY 623 GAAGCTTATGGAATGGGAAGCAGGGATCTCGTCTCTTCTTCAAGTTTCTAGTTTATATGCA 682  
Db 605 AAAGCAGATCTGGCTGGAAATAATGCAACCGAGTCAGTTAAGCTATTTCTTCAACAGCT 664  
QY 683 AATCCAGCTCAACCAAGATTTGTCAGAGAACCAATCTGACCTGTGCTCACTCAACTTATCAT 742  
Db 665 AGTGACATAAACAAGCAATCTGTAGCAAA-----AGGATCAACTAGCAAGCCA 712  
QY 743 CAAATCAAGGGGAAACAATTTCAAGCCCTTTTACGTTGAATTTGCTATGCTTAAACCCCTTATCA 802



Db 713 GCARATAAATCTGAAATCTCCAGAGTCTTTTGAAGGAACCTATGATTCACCTAGCGCC 772  
QY 803 GAAGCCATGTGGAATCTGATGGCTTATTTTCGACCCAGCGCAATCACAAGTCGAACC 862  
Db 773 CAAGTTACAGTGAATCAGATGGCTGTGCTTTGACCTGTAGATTTATGATCGTACA 832  
QY 863 GCCAGAGGTGTAGTGTCCCTCATGTGTAACCATTAACATTCCTTTATGAAACAAATG 922  
Db 833 CCAATGGAGTTGGATTCGCGATGGCGACCATTAACCACTTTATTCCTTACAGCAAGCTT 892  
QY 923 TCTGAATGGAAAACGAATTCGTCGTATTAATTCCTTCCTT 964  
Db 893 TCTGCCTTAGAAGAAAAGATTGCCAGATGGTGCCTATCAGT 934

## RESULT 13

US-09-536-784-181  
; Sequence 181, Application US/09536784  
; Patent No. 6573082  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/536,784

FILING DATE: 30-Oct-1997  
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083

FILING DATE: OCT-30-1997  
; ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks  
; REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB34023  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 181:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 1342 base pairs  
; TYPE: nucleic acid

STRANDEDNESS: double  
; TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 181:  
; US-09-536-784-181

Query Match 17.5%; Score 401; DB 4; Length 1342;  
Best Local Similarity 65.7%; Pred. No. 1.1e-112;  
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

QY 23 CACCAAGTGTGTCAGTTAAGAAAGTCTTAATCGAGTTCTTATATAGATGGTATCAG 82

Db 8 CAGCATCGTTCGCGAGGAATAAGACAATAATCGTCTCTTATGTGGATGGCAGCCAG 67

QY 83 GCTGTCAAAAGCGAGAAAACCTTGACACAGATGAAGTCAGTAAGAGGGAGGGATCAAC 142

Db 68 TCAAGTCAGAAAAGTGAACCTTGACACAGCAGGTTAGCCAGAAAAGAAATTCAG 127

QY 143 GCCGAACAAATNGTGNATCAAGATTACGGATCAAGGTTATGTGACCTCTCATGGAGACCAT 202

Db 128 GGTGAGCAAAATGTATCAAAATTAAGATCAGGGCTATGTAAAGTCACAGGTGACAC 187  
QY 203 TATCATTAATAATGCGCAAGGTTCTTATGATGCCATCATCAGTGAAGAGCTCCTCATG 262  
Db 188 TATCATTAATAATGCGAAAGTTCTTATGATGCCCTCTTATGTAAGAACTCTTGATG 247  
QY 263 AAAGATCCGAATTAATCAGTTGAAGGATTGAGACATTTGATGTAATGAAATCAAGGGTGGTTAT 322  
Db 248 AAGGATCCAAACTATCAACTTAAAGACGCTGATATTGTCAATGAAGTCAAGGGTGGTTAT 307  
QY 323 GTCAATTAAGGTAACCGTAAATTAATGATGTTAAAGATGAGGATGAGGATGAGGATTAAT 382  
Db 308 ATCATCAGGTCGATGGAATAATTAATGTTACTCTGAAAGATGAGGATGAGGATTAAT 367  
QY 383 ATTGCGACAAAAGAGAGATTAAACGTCAGAACGAGAACGAGTCAATAATCAATACTCA 442  
Db 368 GTTCGAACTAAGATGAATCAATCGTCAAAACAAAGAAC---ATGTCAAAGATAATGAG 424  
QY 443 AGAGCAGATAATGCTGTTGTCGAGCCAGAGCCAAAGGAGGTTATACAAAGGATGATGGG 502  
Db 425 AAGGTTAACTCTAAATGTTGCTGTAGCAAGGTCCTCAGGACGATATACGACAAATGATGGT 484  
QY 503 TATATCTTCAATGCACTCTGATATCATGAGGACGAGGATGATGATGATGATGATGATGATGAT 562  
Db 485 TATGTCCTTTAATCCAGCTGATATTAATGAGGATGAGGATGATGATGATGATGATGATGAT 544  
QY 563 GCGGACCATTAACCATTAACATTCCTAAGATGAGTATCAGTACGAGGATGATGATGATGATGATGAT 622  
Db 545 GGAGGTCATCTACTACTACATTTCCAAAGGATTTATCTGCTAGTGAATGAGCAGAGCT 604  
QY 623 GAAGCCTATTGGAATGGGAAGCAGGATCTGTCCTTCTTCAAGTTCTAGTATTAATGCA 682  
Db 605 AAAGCACATCTGGCTGGAATAATATGCAACCGAGTCAGTTAAGCTATTCTTCAACAGCT 664  
QY 683 AATCCAGCTCAACCAAGATTGTGAGAGAACCAACAACTGACTGTCTCACTCAACTTATCAT 742  
Db 665 AGTGACATAACACGCAATCTGTAGCAAA-----AGGATCAACTAGCAAGCCA 712  
QY 743 CAAATCAAGGGGAAAACATTTCAAGCCTTTTACGTAATGATGCTAAACCCCTTATCA 802  
Db 713 GCAATTAATCTGAHAATCTCCAGATCTTTTGAAGAACTCTATGATTCACCTAGCGCC 772  
QY 803 GAACGCCATGTGGAATCTGATGCCCTTATTTTCGACCCAGCGCAAAATCAAAAGTCGAACC 862  
Db 773 CAACGTTACAGTGAATCAGATGCGCTGGTCTTTGACCCCTGCTAAGATTATCAGTCGTACA 832  
QY 863 GCGAGAGGTGTAGTGTCCCTCATGTGTAACCATTAACACTTTATCCCTTATGAAACAAATG 922  
Db 833 CCAAAATGGAGTTGCGATTCGATTCGATGCGGACCATTAACCACTTTATTCCTTACAGCAAGCTT 892  
QY 923 TCTGAATTCGAAAACGAATTCGTCGTATTAATTTCCCTTCGT 964  
Db 893 TCTGCCTTAGAAGAAAAGATTGCCAGATGGTGCCTATCAGT 934

## RESULT 14

US-09-468-656A-7  
; Sequence 7, Application US/09468656A

; Patent No. 6582706  
; GENERAL INFORMATION:

; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.

; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Group A and B Having Selected Structural

; FILE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-444

; CURRENT APPLICATION NUMBER: US/09/468,656A  
; CURRENT FILING DATE: 1999-12-02

; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21

; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 7

```

; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-468-656A-7

Query Match      17.5%; Score 401; DB 4; Length 1455;
Best Local Similarity 65.7%; Pred. No. 1.2e-112;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

20 23 CACCAAGCTGGTTCAGGTTAAGAAAGAGTCTAAATCGAGTTCTTATATAGATGGTGATCAG 82
    |||
79 CAGCATCGTTCGAGGAAATATAGCAATATCTGTCTCTTATGTGGATGGCAGCAG 138
    |||
83 GCTGGTCAAAAGGCAGAAAATCTTGACACCAAGATGAAGTCAAGTACAGAGAGGGGATCAAC 142
    |||
139 TCAAGTCAGAAAAGTGAAAATCTTGACACCAAGCAGGTTAGCCAGAAAGAAAGGATTCAG 198
    |||
143 GCCGAACAATNGTNATCAAGNTACGGATCAAGGTTATGTGACCTCTCATGGAGACCAT 202
    |||
199 GCTGAGCAAAATCTGAATCAAAATTTACAGATCAGGGCTATGTAACTGACACGCTGACCC 258
    |||
203 TATCATTTACTATATATGCAAGGTTCTTATGATGCCATCATCAGTGAAGAGCTCTCATG 262
    |||
259 TATCATTTACTATATGGAAGTTCTTATGATGCCCTCTTTAGTGAAGAACTCTTGATG 318
    |||
263 AAAGATCCGAATTTATCAGTTGAAGGATTCAGACATTTCTCAATGAAATCAAGGGTGGTTAT 322
    |||
319 AAGGATCCAAACTATCAACTTAAAGCGCTGATATTTCTCAATGAAGTCAAGGGTGGTTAT 378
    |||
323 GTCATTAAGGTAACGGTAAATCTATGTNNATCTTAAAGATGAGCTCATGCGGATAAT 382
    |||
379 ATCATCAAGTCTGATGAAAATATTTATGTCTACCTGAAGATGAGCTCATGCTGATAAT 438
    |||
383 ATTCGACAAAAGAGATTAACGTCAAGACGAGAAAGCAGTCAATATCAATCAACTCA 442
    |||
439 GTTCGAAGTAAAGATGAATCAATCTGTCGACAAAACAGAAC---ATGTCGAAGATAATGAG 495
    |||
443 AGAGCAGATAATCTGTGCTGACGACGAGCCAGAGCTTATATCAACCGATGATGGG 502
    |||
496 AAGTTAACTCTAATGTTGCTGACGAGGCTCTCAGGAGCATATACGACAAATGATGGT 555
    |||
503 TATATCTCAATGATCATCTATATCAATTTAGGACACACGGGTGATGTTATGCTTCCTCAC 562
    |||
556 TATGCTCTTAAATCCAGTGTATATATCGAAGATGACGGGTAATGTTATATCGTTCTCAT 615
    |||
563 GCGGACCATTACCATTTACATCTCTAAGATGAGTTATCAGTACGAGTATGCTGCTGCA 622
    |||
616 GGAGGTCACTATCACTACATCTCCAAAAGCGATTTATCTGCTAGTGAATTAGCAGCAGCT 675
    |||
623 GAAGCCTATTGGAATGGGAAGCAGGGATCTCGTCTCTTCAAGTCTAGTTATATGCA 682
    |||
676 AAGACCATCTGCTGGAAAATAATGCAACCGAGTCAGTTAGCTATCTTTCACAGCT 735
    |||
683 AATCCAGCTCAACCAAGATTTGTGAGAGAACCAAAATCTGACTGTCTCACTCAACTTATCAT 742
    |||
736 AGTGACAATAACACGCAATCTGTAGCAA-----AGGATCAACTAGCAAGCCA 783
    |||
743 CAAATCAAGGGGAAAACATTTCAAGCCTTTTACGTTAATGTATGCTAAACCTTATCA 802
    |||
784 GCAAAATAATCTGAAAATCTCCAGAGTCTTTTGAAGGAATCTCTATGATTCACCTAGCGCC 843
    |||
803 GAAGCCATGTGAATCTGATGSCCTTATTTTCGACCCAGCGCAATCAAGTTCGAAACC 862
    |||
844 CACAGTTACAGTGAATCAGATGSCCTGCTTTTGAACCTCTGTAAGATTAATCAGTCGTA 903
    |||
863 GCCAGAGGTGATGCTGCTCCCTCATGTGTAACCATTAACCATTTTATCCCTTATGAACAAATG 922
    |||
904 CCAATGAGAGTTCGATTCGGATGGGACCAATACACCTTATTCCTTACAGCAAGCTT 963
    |||
923 TGTGAATGGAAAACGAATTTGCTCGTATTTATTCCTTCCTT 964
    |||
964 TCTGCCTTAGAAGAAAAGATTGCGAAGATGGTGCCTATCAGT 1005
    |||

Query Match      17.5%; Score 401; DB 4; Length 6867;
Best Local Similarity 65.7%; Pred. No. 3.4e-112;
Matches 619; Conservative 0; Mismatches 308; Indels 15; Gaps 2;

20 23 CACCAAGCTGGTTCAGGTTAAGAAAGAGTCTAAATCGAGTTCTTATATAGATGGTGATCAG 82
    |||
6707 CAGCATCGTTCGAGGAAATATAGGACATATATCGTGTCTCTTATGTGATGGCAGCCAG 6648
    |||
83 GCTGGTCAAAAGGCAGAAAATCTTGACACGAGATGAAGTCAAGTGAAGAGGGGATCAAC 142
    |||
6647 TCAAGTCAGAAAAGTGAAAATCTTGACACCAAGCAGGTTAGCCAGAAAGAAAGGATTCAG 6588
    |||
143 GCCGAACAATNGTNATCAAGNTACGGATCAAGGTTATGTGACCTCTCATGGAGACCAT 202
    |||
6587 GCTGAGCAAAATGTGAATCAAAATTTACAGATCAGGGCTATGTAACTGATCAACCGTGCAC 6528
    |||
203 TATCATTTACTATATGCGCAAGGTTCTTATGATGCCATCATCAGTGAAGAGCTCCTCATG 262
    |||
6527 TATCATTTACTATATGGAAGTTCTTATGATGCCCTCTTTAGTGAAGAACTCTTGTATG 6468
    |||
263 AAAGATCCGAATTTATCAGTTGAAGGATTCAGACATTTGTCAATGAAATCAAGGGTGGTTAT 322
    |||
6467 AAGGATCCAAACTATCAACTTAAAGCGCTGATATTTGTCATGAAGTCAAGGGTGGTTAT 6408
    |||
323 GTCATTAAGGTAACGGTAAATATCTATGTTNATCTTAAAGATGACGCTCATGGGATAAT 382
    |||
6407 ATCATCAAGGTCGATGGAAAATATTTATGTCTACCTGAAAGATGACGCTCATGCTGATAAT 6348
    |||
383 ATTCGACAAAAGAGATTAACGTCAAGAGCAGGAGCAGTCAATATCAATCAATCACTCA 442
    |||
```

Db 6347 GTTCGAACAAAGATGAATCAATCGTCAAAACAGAAC---ATGTCAGAGTAATGAG 6291  
QY 443 AGACGAGATATGCTGTGTCGACCCAGAGACCGGATATACACGAGATGATGG 502  
Db 6290 AAGGTTAACTCTAATGTTGCTAGCAGAGGCTCAGGAGCATATACGACAAATGATGT 6231  
QY 503 TATATCTTAAATGATCATGATATCATGAGACACGGGTGATGCTTATATGTTCTCTAC 562  
Db 6230 TATGCTTTTATCCAGCTGATATATCGAAGATACGGGTAAATGCTTATATGTTCTCTCAT 6171  
QY 563 GGCGACCATPACCATTAATCTTCAAGATCAGTATCAGCTAGGAGTACCTGCTCA 622  
Db 6170 GGAGGTCATCTACTACTATCCCAAAAGCGATTATCTGCTAGTGAATAGCAGCAGCT 6111  
QY 623 GAAGCTTATGGAATGGAGACGAGGATCTGCTCTTCTTCAAGTCTTATGTTATATGCA 682  
Db 6110 AAGCACATCTGGCTGGAAAAAATATGCAACGGAGTCAGTTAAGCTATCTTCAACAGCT 6051  
QY 683 AATCCAGCTCAACCAAGATTGTCAGAGAACCAACAATCTGACTGTCACTCCAACTTATCAT 742  
Db 6050 AGTGACAAATACACGCAATCTGTAGCAAA-----AGGATCAACTAGCAAGCCA 6003  
QY 743 CAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGAATGTTATGCTTAAACCTTATCA 802  
Db 6002 GCAATTAATCTGAAATCTCCAGAGTCTTTTGAAGGAATCTATGATTCACCTAGCGCC 5943  
QY 803 GAAGCCATGTGGAATCTGATGGCTTATTTTGAACCCAGCGCAATCACAAGTCGAACC 862  
Db 5942 CAAGTTACAGTGAATCAGATGGCTGCTCTTTGACCTCTGATGATTCAGTCTGACA 5883  
QY 863 GCACAGGTGATGCTGCTCCCTGATGTTAACCATTACCACTTTATCCCTTATGAACAAATG 922  
Db 5882 CCAATAGGATGCGATTCCGATGGCGACCAATACCACTTTATTCCTTACAGCAAGCTT 5823  
QY 923 TCTCAATTTGAAAGAAAGATGCTGCTGATTTATTCCTTCTG 964  
Db 5822 TCTGCTTAGAAGAAAGATTTGCCAGATGCTGCTATCAT 5781

RESULT 16  
US-08-743-637B-34/c  
; Sequence 34, Application US/08743637B  
; Patent No. 5994066  
; GENERAL INFORMATION:  
; APPLICANT: BERGERON, Michel G.  
; APPLICANT: PICARD, Francois J.  
; APPLICANT: OUELLETTE, Marc  
; APPLICANT: ROY, Paul H.  
; TITLE OF INVENTION: SPECIES-SPECIFIC AND UNIVERSAL DNA  
; TITLE OF INVENTION: PROBES AND AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND  
; TITLE OF INVENTION: IDENTIFY COMMON BACTERIAL PATHOGENS AND ASSOCIATED  
; TITLE OF INVENTION: ANTIBIOTIC RESISTANCE GENES FROM CLINICAL SPECIMENS ...  
; NUMBER OF SEQUENCES: 273  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: QUARLES & BRADY  
; STREET: 411 EAST WISCONSIN AVENUE  
; CITY: MILWAUKEE  
; STATE: WISCONSIN  
; COUNTRY: USA  
; ZIP: 53202-4497  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/743,637B  
; FILING DATE: 04-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/526,840  
; FILING DATE: 11-SEP-1995  
; ATTORNEY/AGENT INFORMATION:

; NAME: BAKER, Jean C.  
; REGISTRATION NUMBER: 35,433  
; REFERENCE/DOCKET NUMBER: 850586.90012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (414) 277-5000  
; TELEFAX: (414) 277-5591  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 841 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pneumoniae  
; US-08-743-637B-34  
  
Query Match 11.9%; Score 272.4; DB 2; Length 841;  
Best Local Similarity 92.9%; Pred. No. 2.7e-73;  
Matches 325; Conservative 0; Mismatches 6; Indels 19; Gaps 3;  
  
QY 1956 TTTTGGCGACTGTCAAGTACTATGTCGAACATCCAAACGAACTCCGCAATTCAGATAATG 2015  
Db 841 TTTTGGCGACTGTCAAGTACTATGTCGAACATCCAAACGAACTCCGCAATTCAGATAATG 783  
QY 2016 GTTTTGGTAACGCTAGCGACCATGTTCAAAGAAACAAATAATGGTCAAGCTGATACCAATC 2075  
Db 782 GTTTTGGTAACGCTAGCGACCATGTTCAAAGAAACAAATAATGGTCAAGCTGATACCAATC 723  
QY 2076 AAACGAAACAAACGAGGAGGAAACCTCAGACAGAAACCTGAGGAGAAACCCCTC 2135  
Db 722 AAACGAAACAAACGAGGAGGAAACCTCAGACAGAAACCTGAGGAGAAACCCCTC 663  
QY 2136 GAGAGAGAAACCGCAAGCGAGAACACAGAGTCTCCAAACCAACAGAGGA----- 2187  
Db 662 GAGAGAGAAACCGCAAGCGAGAACACAGAGTCTCCAAACCAACAGAGGAACCAAG 603  
QY 2188 -----ACCAGAAGAAATCACCAGAGGAATCAGAGAACTCAGAGTCCGAGACTGAAAGG 2240  
Db 602 AAGAATCACCAGAGAAATCACCAGAGGAATCAGAGAACTCAGAGTCCGAGACTGAAAGG 543  
QY 2241 TTGAAGAAACCTGAGAGAGGCTGAAAGATTTACTTGGAAAAATCCAGGAT 2290  
Db 542 T---AAGAAACTGAGAGAGGCTGAAAGATTTACTTGGAAAAATCCAGAA 496

## RESULT 17

US-08-526-840B-34/c  
; Sequence 34, Application US/08526840B  
; Patent No. 6001564  
; GENERAL INFORMATION:  
; APPLICANT: BERGERON, Michel G.  
; APPLICANT: OUELLETTE, Marc  
; APPLICANT: ROY, Paul H.  
; TITLE OF INVENTION: SPECIFIC AND UNIVERSAL PROBES AND  
; TITLE OF INVENTION: AMPLIFICATION PRIMERS TO RAPIDLY DETECT AND IDENTIFY  
; TITLE OF INVENTION: COMMON BACTERIAL PATHOGENS AND ANTIBIOTIC RESISTANCE GENES  
; TITLE OF INVENTION: FROM CLINICAL SPECIMENS FOR ROUTINE DIAGNOSIS IN ...  
; NUMBER OF SEQUENCES: 177  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: QUARLES & BRADY  
; STREET: 411 East Wisconsin Avenue  
; CITY: Milwaukee  
; STATE: Wisconsin  
; COUNTRY: USA  
; ZIP: 53202-4497  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/526,840B

;; FILING DATE: 11-SEP-1995  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/304,732  
;; FILING DATE: 12-SEP-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BAKER, Jean C.  
;; REGISTRATION NUMBER: 35,433  
;; REFERENCE/DOCKET NUMBER: 850586.90012  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (414) 277-5000  
;; TELEFAX: (414) 277-5591  
;; INFORMATION FOR SEQ ID NO: 34:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 841 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; ORIGINAL SOURCE:  
;; ORGANISM: Streptococcus pneumoniae  
;; US-08-526-840B-34

Query Match 11.9%; Score 272.4; DB 3; Length 841;  
Best Local Similarity 92.9%; Pred.No. 2.7e-73;  
Matches 325; Conservative 0; Mismatches 6; Indels 19; Gaps 3;  
QY 1956 TTTTGGCGAGTCTCAAGTACTATGTGCAATCAACGAAAGTCCCGATTCAGATTAATG 2015  
DB 841 TTTTGGCGAGTCTCAAGTACTATGTGCAATCAACGAAAGTCCCGATTCAGATTAATG 783  
QY 2016 GTTTTGGTAACGCTAGCGACCATGTTCAAGAAACAAATGTTCAAGCTGATACCAATC 2075  
DB 782 GTTTTGGTAACGCTAGCGACCATGTTCAAGAAACAAATGTTCAAGCTGATACCAATC 723  
QY 2076 AAACGGAAAAACCAACGAGGAGAAACCTCAGACAGAAAAACCTTGAGGAAGAAACCCCTC 2135  
DB 722 AAACGGAAAAACCAACGAGGAGAAACCTCAGACAGAAAAACCTTGAGGAAGAAACCCCTC 663  
QY 2136 GAGAGAGAAACCGGAAAGGAGAAACAGAGTCTCCAAAAACCAACAGAGGA 2187  
DB 662 GAGAGAGAAACCGGAAAGGAGAAACAGAGTCTCCAAAAACCAACAGAGGA 603  
QY 2188 -----ACCAGAAGATCACAGAGGAATCAGAAGAACTCAGGTGGAGCTGAAAGG 2240  
DB 602 AAGATCACCAAGAAATCACAGAGGAATCAGAAGAACTCAGGTGGAGCTGAAAGG 543  
QY 2241 TTGAGAAAAAATCTGAGAGGCTGAAGATTACTTTGAAAAATCCAGGAT 2290  
DB 542 T-----AAGAAATCTGAGAGGCTGAAGATTACTTTGAAAAATCCAGAT 496

RESULT 18  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/232,463  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/935,313  
;; FILING DATE:  
;; APPLICATION NUMBER: EP 91 114 300.6  
;; FILING DATE: 26-AUG-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BENT, Stephen A.  
;; REGISTRATION NUMBER: 29,768  
;; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 836-9300  
;; TELEFAX: (703) 683-4109  
;; TELEX: 899149  
;; INFORMATION FOR SEQ ID NO: 14:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 7218 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; IMMEDIATE SOURCE:  
;; CLONE: pTZgpt-Fls  
;; US-08-232-463-14

Query Match 2.8%; Score 64.4; DB 1; Length 7218;  
Best Local Similarity 3.9%; Pred.No. 1e-08;  
Matches 11; Conservative 180; Mismatches 91; Indels 0; Gaps 0;  
QY 2008 AGATAATGTTTGGTAACGCTAGGACCAATGTTCAAGAAACAAATGGTCAAGCTGA 2067  
DB 1453 AGATAGAAGAAATTTGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1394  
QY 2068 TACCAATCAACGGAACCAAGCGAGAGAAACCTCAGACAGAAAAACCTGAGGAAGA 2127  
DB 1393 RRR 1334  
QY 2128 AACCCCTCGAGAAGAGAAACCGCAAGCGAGAGAAACAGAGTCTCCAAAAACCAACAGAGA 2187  
DB 1333 RRR 1274  
QY 2188 ACCAGAGAATCACAGAGGAATCAGAAGAACTCAGGTCCGAGACTGAAAAGTTGAAGA 2247  
DB 1273 RRR 1214  
QY 2248 AAAACTGAGAGAGCTGAAGATTACTTTGAAAAATCCAGGA 2289  
DB 1213 RRR 1172

RESULT 19  
US-08-257-073-4  
; Sequence 4, Application US/08257073  
; Patent No. 5766597  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; APPLICANT: de Taisne, Charles  
; APPLICANT: Tine, John A.  
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE  
; NUMBER OF SEQUENCES: 143  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford, P.C.  
; STREET: 530 Fifth Avenue, 25th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/257,073  
 FILING DATE: 09-JUN-1994  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/075,783  
 FILING DATE: 11-JUN-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/852,305  
 FILING DATE: 18-MAR-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/672,183  
 FILING DATE: 20-MAR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Frommer, William S.  
 REGISTRATION NUMBER: 25,506  
 REFERENCE/DOCKET NUMBER: 454310-2570  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 840-3333  
 TELEFAX: (212) 840-0712  
 TELEX: 425066 CUETMS  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2223 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-257-073-4

Query Match 2.5%; Score 56.2; DB 1; Length 2223;  
 Best Local Similarity 48.0%; Pred. No. 1.5e-06;  
 Matches 160; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 1957 TTGGGCACTGCTCAAGTACTATGTCGAACATCCAAACGAACTCCGCGATTTCAGATAATGG 2016  
 DB 1863 TTATCAAAATGGAACCAAAAATATTTCAACAAATATTTCTAGAAAATGA 1922  
 QY 2017 TTTTGGTAACGCTAGCGACCATGTTCAAGAAACAAATGTCAGCTGTACCAATCA 2076  
 DB 1923 TGTTCTTAATCAAGAAACGGAGGAAGAAATGGAATAACAGTTGAAGCAATCACCAGCA 1982  
 QY 2077 AACGGAAAAACCAACCGGAGGAGAACTTCAGACAGAAAAACCTGAGGAGAAACCCCTCG 2136  
 DB 1983 AATAGAAGCTGAAGTGGATGCCCTGCACCAAAAAATTAAGGAGAGAGAAAGAAAA 2042  
 QY 2137 AGAAGAAACCGCAACCGGAGAAACGAGTCTTCCAAAAACCAACAGAGGAAACCGAGAGA 2196  
 DB 2043 AGAAAAAGAAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 2102  
 QY 2197 ATCAACAGAGGAATCAGAGAACCTCAGGTGAGACTGAAAGGTTGAAGAAAAACTGAG 2256  
 DB 2103 AAAGAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 2162  
 QY 2257 AGAGGCTGAAGATTACTTTGAAAAATCCAGGA 2289  
 DB 2163 ACAGAGAGAGAGAGAGAGAGAGAAATAGTACCAGA 2195

RESULT 20  
 US-09-640-419C-6  
 Sequence 6, Application US/09640419C  
 Patent No. 6630615  
 GENERAL INFORMATION:  
 APPLICANT: Bidney, Dennis L.  
 APPLICANT: Crasta, Oswald R.  
 APPLICANT: Hu, Xu  
 APPLICANT: Lu, Guihua  
 TITLE OF INVENTION: DEFENSE-RELATED SIGNALING GENES AND METHODS OF USE  
 FILE REFERENCE: 35718/199009 (5718-92)  
 CURRENT APPLICATION NUMBER: US/09/640,419C  
 CURRENT FILING DATE: 2000-08-17

PRIOR APPLICATION NUMBER: 60/149,656  
 PRIOR FILING DATE: 1999-08-18  
 PRIOR APPLICATION NUMBER: 60/206,405  
 PRIOR FILING DATE: 2000-05-23  
 NUMBER OF SEQ ID NOS: 28  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 6  
 LENGTH: 865  
 TYPE: DNA  
 ORGANISM: Helianthus annuus  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)..(519)  
 US-09-640-419C-6

Query Match 2.2%; Score 51.4; DB 4; Length 865;  
 Best Local Similarity 50.2%; Pred. No. 2.3e-05;  
 Matches 127; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 2002 GCATTCAGATAATGGTTTGTAACTAGCGATGTTCAAGAAACCAAAAATGTCATCA 2061  
 DB 30 GCATGATGAGAAATTCAGAGGAATCGAATGTCGAATCGAGAGGACATTAAGAGCAC 89  
 QY 2062 AGCTGATACCAATCAAAACGAAAAACCAACGAGGAGAAACCTCAGACAGAAAACTGA 2121  
 DB 90 AATGATCAAGAAACCAACCAACCAACCAAGAGACCGCCCTTAAGAAAAACCTCT 149  
 QY 2122 GGAAGAAACCTTCGAGAGAGAGAAACGGAACCGGAAACCGGAAACCGAGTCTCCAAACCAAC 2181  
 DB 150 TGATGACATAACTGTTGCACCTGAACTGATGTAGTCTCTCTGAACTCTAGAGGCAA 209  
 QY 2182 AGAGGAAACAGAGAAATCAACGAGGAAATCAGAGGAAATCAGAGAAACCTCAGTCTGAGACTGAAAAGGT 2241  
 DB 210 ACCAGAACCAAGGCGCAACCAAGAACCAAGAACCGGAAAGCCCTTGAGCTGAGGT 269  
 QY 2242 TGAAGAAAAACTG 2254  
 DB 270 TGAACCAAAAGTG 282

RESULT 21  
 US-08-956-171E-59  
 Sequence 59, Application US/08956171E  
 Patent No. 6593114  
 GENERAL INFORMATION:  
 APPLICANT: Charles Kunsch  
 Gil H. Choi  
 Patrick S. Dillon  
 Craig A. Rosen  
 Steven C. Barash  
 Michael R. Fannon  
 TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
 NUMBER OF SEQUENCES: 5256  
 CORRESPONDENCE ADDRESS:  
 ADDRESSER: Human Genome Sciences, Inc.  
 STREET: 9410 Key West Avenue  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: USA  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/956,171E  
 FILING DATE: 20-Oct-1997  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/009,861  
 FILING DATE: January 5, 1996  
 APPLICATION NUMBER: 08/781,986

Db	492	TAAGAGCAAGTGGCTACGGATACACAGGCTGTGTGATGAACACACAAAAAGTAGTGGATCA	551
Qy	2008	AGATAATGGTTTGTGAAGCTAGCGACCATGTTCAAAGAAACAAATAATGGTCAAGCTGA	2067
Db	552	AGCTCAACAGATGTTAAACCAACAACAGCTGTGTGTGATGAAAAAGCAAAAGAAACGAA	611
Qy	2068	TACCAATCAAAACGGAAAAACCAAGCGAGGAGAAACCTCGACACAGAAAAACCTGAGGAGA	2127
Db	612	TGCTGCTAAAGTGCACAAATGAAAAAGATCAACAAGCAGTAAACAGCTCGGAAACAGACACA	671
Qy	2128	AACCCCTCGAAGAGAAACCCCAAGCGAGAAACCAAGACTCTCCAAAAACCAACAGAGA	2187
Db	672	AGCCAAGCTTGAAGAATTAGCGAAAAATCGGAAGCGGAAAAAGCAAAAGGCAGAAAAAGA	731
Qy	2188	ACCAGAAGAATCACAGAGGAATCAGAAGAACCTCAGGTCGAGACTGAAAAGGTTGAGA	2247
Db	732	ACAAGCAGCAAAAGAGCTGAACTGGCTTAACAAACAGAAAGAGAAGCCAAAGCTAAAGA	791
Qy	2248	AAAACTGAGAGAGGCTGAAGATTTACTTTGGAAA	2280
Db	792	TCAAAAGCAAGGACGATCAGCAGTCGAGA	824
RESULT 23			
US-09-627-122-21/c			
; Sequence 21, Application US/09627122			
; Patent No. 6472521			
; GENERAL INFORMATION:			
; APPLICANT: Uhlmann, Eugen			
; APPLICANT: Greiner, Beate			
; APPLICANT: Unger, Eberhard			
; APPLICANT: Gothe, Gislinde			
; APPLICANT: Schwerdel, Marc			
; TITLE OF INVENTION: OLIGONUCLEOTIDES FOR THE INHIBITION OF HUMAN eg5			
; TITLE OF INVENTION: EXPRESSION			
; FILE REFERENCE: 02481.1678			
; CURRENT APPLICATION NUMBER: US/09/627,122			
; CURRENT FILING DATE: 2000-07-27			
; NUMBER OF SEQ ID NOS: 23			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 21			
; LENGTH: 5340			
; TYPE: DNA			
; ORGANISM: Plasmodium falciparum			
US-09-627-122-21			
Query Match 2.0%; Score 45.4; DB 4; Length 5340;			
Best Local Similarity 44.3%; Pred. No. 0.0056;			
Matches 178; Conservative 0; Mismatches 224; Indels 0; Gaps 0;			
Qy	9	ATCAACTTGGTTCGTACCAAGCTGGTCAAGTTAAGAAAGAGTCTAATCGAGTTCTTTATA	68
Db	928	ATAATAATTATCTTTATATATGATGATAATAAATGGATCAAGGTATGATCATATAATAAT	869
Qy	69	TAGATGGTGTATCAGGCTGGTCAAAAGGCAGAAAACTTGACACACAGATGAAGTCAGTAAGA	128
Db	868	ATCTTTATATATGATGATAATAAATGGATCAAGGTATGATCATATAATAATTATCTTTATA	809
Qy	129	GGAGGGGATCAACGCCGAAACAAATNGTNAFTCAAGATTACGGATCAAGGTTATGTGACCT	188
Db	808	ATGATGATAATAAAATGGATCAAGGTATCATATAATTTATTTATCTTTATAATGATGATA	749
Qy	189	CTCATGGAGACATTATCATCTATATATGCGAAGGTTCTTTATGATGCCATCATCAGTG	248
Db	748	ATAAATATGATCAAGGTGATCATATAATTATTTATCTTTATAATGATGATATCATTAAGAC	689
Qy	249	AAGAGCTCTCATGAAGATCCGAATTATCAGTTGAAGGATTCAGACATTTGCAATGAAA	308
Db	688	AAAAATTATCTTGGAGTTTAAAAATACCTTGATGAAGAGATAGAACAAATTACACAA	629
Qy	309	TCAAGGTTGGTTATGTCAATTAAGGTAAACGGTAAATACTATGTTNATACCTTTAAGGATCGAG	368
Db	628	CACATGATATAGATTTAAGTAAAGTATTTTGAACAAATCGATGAACAACTTCATTTCTTTA	569



Patent No. 6231861  
GENERAL INFORMATION:  
APPLICANT: Barnwell, John  
TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,  
TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby and Darby  
STREET: 805 Third Ave.  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022-7513  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/092,458  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/719,821  
FILING DATE: 09/30/96  
ATTORNEY/AGENT INFORMATION:  
NAME: Gogoris, Adda  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17686US3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3337 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmodium vivax  
IMMEDIATE SOURCE:  
CLONE: PYMB3.3.1  
US-09-092-458-1

Query Match 2.0%; Score 44.8; DB 3; Length 3337;  
Best Local Similarity 50.0%; Pred. No. 0.0062;  
Matches 112; Conservative 0; Mismatches 112; Indels 0; Gaps 0;  
QY 2066 GATACCAATCAACGGAACCAACGAGGAGGAGAACTTCAGACAGAGAAACCTTGAGGAA 2125  
Db 2070 GAACTCGACAGAGGAGAGAGAGTGGAGAGGTACCTCGACAGAGTAGAAGTGGAGAG 2129  
QY 2126 GAAACCTTCGAGAGAGAGAACCGAAGCGAGAGAACAGAGTCTCCAAACCAACAGAG 2185  
Db 2130 GTACCTCGAGAGTAGAAGAGAGTGGAGAGGTACCAAGAGAGTAGAAGAGGTACCCGCA 2189  
QY 2186 GAACCAAGAGATCAACCAAGAGATCAGAGAGACCTCAGTTCGAGACTGAAAGGTTGAA 2245  
Db 2190 GAAGTAGAAGAGTGGAGAGGTACCAAGAGAGTGGAGAGGTACCAAGAGAGTGGAA 2249  
QY 2246 GAAAACTCAGAGAGGTGAGAGTTTACTTGGAAAAATCCAGGA 2289  
Db 2250 GAGTACCAGAGAGTGGAGAGGTACCAAGAGAGTGGAGAG 2293

RESULT 27  
US-08-742-185-101  
; Sequence 101, Application US/08742185  
; Patent No. 6020476

GENERAL INFORMATION:  
APPLICANT: Page, David C.  
APPLICANT: Reijo, Renee  
APPLICANT: Saxena, Richa  
APPLICANT: Hawkins, Trevor  
APPLICANT: Reeve, Mary Pat  
TITLE OF INVENTION: DAZ: A GENE FAMILY ASSOCIATED WITH AZOOSPERMIA  
NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/742,185  
FILING DATE: 30-OCT-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/690,734  
FILING DATE: 31-JUL-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/310,429  
FILING DATE: 22-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: WHI94-07A2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540  
INFORMATION FOR SEQ ID NO: 101:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43795 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-742-185-101

Query Match 1.9%; Score 44.2; DB 3; Length 43795;  
Best Local Similarity 49.4%; Pred. No. 0.056;  
Matches 115; Conservative 0; Mismatches 118; Indels 0; Gaps 0;  
QY 2043 AAGAAACAAATAATGGTCAAGCTGATACCAATCAACCGGAAACCAACGAGGAGAAAC 2102  
Db 38825 AAAAAAAAAAAAAAAAAAAAAAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 38884  
QY 2103 CTCAGACAGAAACCTCGAGGAGAAACCCCTCGAGAGAGAAACCGCAACGAGGAGAAAC 2162  
Db 38885 GAGGAGGAGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38944  
QY 2163 CAGAGTCTCCAAACCAACAGAGAGAACAGAGAGATCACCAGAGAGATCAGAGAGACCTC 2222  
Db 38945 AAG 39004  
QY 2223 AGGTGAGAGCTGAAAGAGTTGAAAGAGAACTGAGAGAGGCTGAAAGATTACTT 2275  
Db 39005 CAGAGAGAGAGAGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 39057

RESULT 28  
US-08-914-479A-3  
; Sequence 3, Application US/08914479A  
; Patent No. 6419932  
; GENERAL INFORMATION:  
; APPLICANT: Dale, James B.



;; TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER  
;; FILE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE  
;; FILE REFERENCE: 481112.404C2  
;; CURRENT APPLICATION NUMBER: US/08/914,479A  
;; CURRENT FILING DATE: 1997-08-19  
;; PRIOR APPLICATION NUMBER: 08/409,270  
;; PRIOR FILING DATE: 1995-03-23  
;; PRIOR APPLICATION NUMBER: 07/945,860  
;; PRIOR FILING DATE: 1992-09-16  
;; NUMBER OF SEQ ID NOS: 19  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 3  
;; LENGTH: 765  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: An antigen of M5 and a carrier of the  
;; OTHER INFORMATION: COOH-terminal portion of M5  
US-08-914-479A-3

Query Match 1.9%; Score 43.8; DB 4; Length 765;  
Best Local Similarity 48.2%; Pred. No. 0.0045;  
Matches 123; Conservative 0; Mismatches 132; Indels 0; Gaps 0;  
QY 1961 GCGACTGTCAAGTACTATGTGCAACATCCAAACGATCCGCGATTTCAGATAATGTTTT 2020  
DB 280 GCAAGTCGCAAGGCTTCGCGTGAATTAGACGATCAGTGAAGCTAAGAAACAAAGTT 339  
QY 2021 GGTAAAGCTAGCGACCATGTTCAAGAAACAAATAATGTTCAAGCTGATACCAATCAAACG 2080  
DB 340 GAAAGAGCTTTAGAGAGCAAGCAACAGCAATAGTCTCTTTGAAAGAACTTAACAAAGAG 399  
QY 2081 GAAAGAACCAAGCGAGAGAAACCTCAGACAGAAAAACCTGAGGAAGAACCCCTCGAGAA 2140  
DB 400 CTTGAGAGAAAGCAAGAAATTAACAGAAAGAAAGAAAGCTGACTACAGCAAACTTGA 459  
QY 2141 GAGAAACCGAAAGCGAGAAACCGAGAGTCTCCAAACCAACAGAGAAACCAAGAGAAATCA 2200  
DB 460 GCAGAGCAAAAGCACTCAAGAGAACTAGCAAAACAGCTGAAGAACTTGCAAAACCTA 519  
QY 2201 CCAGAGGAATCAGAA 2215  
DB 520 AGAGCTGGAAAGCA 534

RESULT 29  
US-08-914-479A-5  
; Sequence 5, Application US/08914479A  
; Patent No. 6419932  
; GENERAL INFORMATION:  
; APPLICANT: Dale, James B.  
; TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER  
; FILE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE  
; FILE REFERENCE: 481112.404C2  
; CURRENT APPLICATION NUMBER: US/08/914,479A  
; CURRENT FILING DATE: 1997-08-19  
; PRIOR APPLICATION NUMBER: 08/409,270  
; PRIOR FILING DATE: 1995-03-23  
; PRIOR APPLICATION NUMBER: 07/945,860  
; PRIOR FILING DATE: 1992-09-16  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 855  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: An antigen of three segments of M5 and a carrier  
; OTHER INFORMATION: of the COOH-terminal portion of M5  
US-08-914-479A-5

Query Match 1.9%; Score 43.8; DB 4; Length 855;  
Best Local Similarity 48.2%; Pred. No. 0.0049;

Matches 123; Conservative 0; Mismatches 132; Indels 0; Gaps 0;  
QY 1961 GCGACTGTCAAGTACTATGTGCAACATCCAAACGATCCGCGATTTCAGATAATGTTTT 2020  
DB 370 GCAAGTCGCAAGGCTTCGCGTGAATTAGACGATCAGTGAAGCTAAGAAACAAAGTT 429  
QY 2021 GGTAAAGCTAGCGACCATGTTCAAGAAACAAATAATGTTCAAGCTGATACCAATCAAACG 2080  
DB 430 GAAAGAGCTTTAGAGAGCAAGCAACAGCAATAGTCTCTTTGAAAGAACTTAACAAAGAG 489  
QY 2081 GAAAGAACCAAGCGAGAGAAACCTCAGACAGAAAAACCTGAGGAAGAACCCCTCGAGAA 2140  
DB 490 CTTGAGAGAAAGCAAGAAATTAACAGAAAGAAAGCTGAGCTACAGCAAACTTGA 549  
QY 2141 GAGAAACCGAAAGCGAGAAACCGAGAGTCTCCAAACCAACAGAGAAACCAAGAGAAATCA 2200  
DB 550 GCAGAGCAAAAGCACTCAAGAGAACTAGCAAAACAGCTGAAGAACTTGCAAAACCTA 609  
QY 2201 CCAGAGGAATCAGAA 2215  
DB 610 AGAGCTGGAAAGCA 624

RESULT 30  
US-08-937-271-9  
; Sequence 9, Application US/08937271  
; Patent No. 6063386  
; GENERAL INFORMATION:  
; APPLICANT: Dale, James B.  
; APPLICANT: Lederer, James W.  
; TITLE OF INVENTION: RECOMBINANT MULTIVALENT M PROTEIN  
; TITLE OF INVENTION: VACCINE  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/937,271  
; FILING DATE: 15-SEP-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Roseman, Stephen J.  
; REGISTRATION NUMBER: 43,058  
; REFERENCE/DOCKET NUMBER: 481112.405C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 918 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..915  
US-08-937-271-9

Query Match 1.9%; Score 43.8; DB 3; Length 918;

Best Local Similarity 48.2%; Pred. No. 0.0052;  
Matches 123; Conservative 0; Mismatches 132; Indels 0; Gaps 0;  
QY 1961 GCGACTGTCAAGTACTATCTCGACATCCAAACGAAGTCGCGATTCAGATATGTTTT 2020  
Db 433 GCAAGTCGAAAAGCGCTTCGCGGTGATTTAGACGCATCAGCTGAAGCTAAGAACAAAGTT 492  
QY 2021 GGTAAAGCTTAGCGACCATCTTCAAGAAAACAAAAATGGTCAAGCTGATCAATCAACG 2080  
Db 493 GAAAGCTTTAGAGAGCAACACAGCAATATTAGTCTTTGAAACCTTACAAAGAG 552  
QY 2081 GAAACCAAGCGAGAGAAACCTTCAGACAGAAAAACCTCGAGAGAAACCCCTCGAGAA 2140  
Db 553 CTTGAGAAAGCAAGAAATTAACAGAAAAAGAAAGCTGAGCTACAAGCAAACTTGAA 612  
QY 2141 GAGAACCCGAAGCGAGAAACCAAGAGTCTTCAAAACCAACAGAGAACAGAGATCA 2200  
Db 613 CGAGAGCAAAAGCACTCAAGAACCAATATTAGCAAAACAAAGCTGAAGAACTTGCAAACTA 672  
QY 2201 CCAGAGGAATCAGAA 2215  
Db 673 AGAGCTGGAAGCA 687

RESULT 31  
US-09-765-272a-65.rni  
; Patent No. 5231168  
; APPLICANT: DZIEGIEL, MORTEN; BORRE, MARTIN; JEPSEN, SOREN;  
; VUUST, JENS; RIENECK, KLAUS; WIND, ANNETTE; JAKOBSEN, PALLE H.  
; TITLE OF INVENTION: MALARIA ANTIGEN  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION NUMBER: 18-SEP-1989  
; FILING DATE: 18-SEP-1989  
; SEQ ID NO: 1:  
; LENGTH: 3095  
5231168-1  
Query Match 1.9%; Score 43.2; DB 6; Length 3095;  
Best Local Similarity 52.8%; Pred. No. 0.018;  
Matches 93; Conservative 0; Mismatches 83; Indels 0; Gaps 0;  
QY 2094 AGGAGAACTCTGAGACAGAAAAACCTCGAGAGAAACCCCTCGAGAGAGAGAAACCGCAAA 2153  
Db 1244 ATGAAAAGGTCAACATGAATAGTAGAGGTGAGAAATCTACCAAGATGATATAA 1303  
QY 2154 GCAGAAACAGAGTCTTCAAAACCAACAGAGAGAACAGAGAAATCAACAGAGGAATCAG 2213  
Db 1304 ATGAAAAGTTGAACATGAATAGTAGAGTTGAAGAAATCTACCAAGAGATATAAATG 1363  
QY 2214 AGAACCTGAGTCTGAGACTGAAAGGTTGAAGAAACCTGAGAGAGGCTGAAGAT 2269  
Db 1364 AAAAGGTCAACATGAATAGTAGAGTTGAAGAAATCTACCAAGAGATATAAAT 1419

RESULT 32  
US-09-621-976-10407  
; Sequence 10407, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 10407  
; LENGTH: 478  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-621-976-10407  
Query Match 1.9%; Score 43; DB 4; Length 478;  
Best Local Similarity 51.9%; Pred. No. 0.0058;  
Matches 122; Conservative 0; Mismatches 110; Indels 3; Gaps 1;  
QY 2046 GAAACAAAAATGGTCAAGCTGATACCAATCAACCGAAAAACCAAGCGAGGAGAAACCTC 2105  
Db 7 GAAAGGAAATGGAGAGATGGAAGAGAGAAAGGAGAGATGAAAAAGAGGAGAGAGACA 66  
QY 2106 AGACAGAAAAACCTGAGAGAGAAACCCCTCGAGAGAGAGAAACCGAAACGAGAGAACAG 2165  
Db 67 GAAAGGAAACAGGAGATGGAAGAGAGAAATGAGATGGAAGAGAGAGAGGAGAGATATAA 126  
QY 2166 AGTCTCCAAACCAACAGAGAGAACCGAGAGATCAACGAGGAGATCAAGAGAACCTCAGG 2225  
Db 127 AGGGGAAAGATGTAAGAGTCAAGAGAGAGATGAAAA--AGAGAGAGAGATGGAAGAGAG 183  
QY 2226 TCGAGACTGAAAAAGGTTGAAGAAAACTGAGAGAGGCTGAAGATTTACTTTGAAA 2280  
Db 184 ATGAGGTGGAATGAGAGAGAGAGCTGGAAAGAGAGAGATTTAAAGAGAGA 238

RESULT 33  
US-09-286-981B-24  
; Sequence 24, Application US/09286981B  
; Patent No. 6503511  
; GENERAL INFORMATION:  
; APPLICANT: Wisemann, Theresa M.  
; APPLICANT: Koenig, Scott  
; APPLICANT: Johnson, Leslie S  
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines  
; FILE REFERENCE: 469201-396  
; CURRENT APPLICATION NUMBER: US/09/286,981B  
; CURRENT FILING DATE: 1999-04-06  
; PRIOR APPLICATION NUMBER: US 60/085,743  
; FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 24  
; LENGTH: 1284  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: cDNA derived  
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae  
US-09-286-981B-24

Query Match 1.9%; Score 42.6; DB 4; Length 1284;  
Best Local Similarity 47.3%; Pred. No. 0.015;  
Matches 129; Conservative 0; Mismatches 144; Indels 0; Gaps 0;  
QY 2009 GATATGTTTTTGGTAAACGCTAGCGACCATGTTTCAAGAAACAAATAATGGTCAAGCTGAT 2068  
Db 913 GAACITGAAATTTGCTGAGTCCGATCGGAAGTTAAAAAGCGAGCTTGAACCTAGTAAAA 972  
QY 2069 ACCAATCAACCGAAAAACCAAGCGAGAGAAACCTCAGACAGAGAAACCTGAGGAGAA 2128  
Db 973 GAGGAAGCTTAAGGAACCTCGAAACGAGGAAAAAGTTAAGCAAGCAAAAGCGAAGTTGAG 1032  
QY 2129 ACCCTCGAGAGAGAAACCGCAAGCGAGAGAACCGAGAGTCTCCAAACCAACAGAGGA 2188  
Db 1033 AGTAAACAAAGCTGAGGCTAAGGTTAGAAAAACATCAAGACAGAGATCGTAAAAAGCAGAA 1092  
QY 2189 CCAGAGAATCACCAGAGGAATCAGAAGAACTCAGGTGCGAGACTGAAAGGTTGAGAA 2248  
Db 1093 GAAGAGCTTAAACGAAAGCAGCAGAGAGATAGTTAAGAAAAACCAAGCTGAGCAA 1152  
QY 2249 AAACGTGAGAGGCTGAAGATTTACTTTGAAA 2281  
Db 1153 CCACAAACGAGCGCGCTCCTCAACAGAGAAA 1185

## RESULT 34

US-09-671-317-14/c  
; Sequence 14, Application US/09671317  
; Patent No. 6528260  
; GENERAL INFORMATION:  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya  
; APPLICANT: Bougueleret, Lydie  
; APPLICANT: Cohen, Annick  
; TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM  
; FILE REFERENCE: 62 US3 CIP  
; CURRENT APPLICATION NUMBER: US/09/671,317  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US 09/536,178  
; PRIOR FILING DATE: 2000-03-23  
; PRIOR APPLICATION NUMBER: PCT/IB00/00403  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: US 60/126,269  
; PRIOR FILING DATE: 1999-03-25  
; PRIOR APPLICATION NUMBER: US 60/131,961  
; PRIOR FILING DATE: 1999-04-30  
; NUMBER OF SEQ ID NOS: 977  
; SOFTWARE: Patent.pm  
; SEQ ID NO 14  
; LENGTH: 929  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 501  
; OTHER INFORMATION: 12-454-363 : polymorphic base A or G  
; NAME/KEY: misc.binding  
; LOCATION: 481..500  
; OTHER INFORMATION: 12-454-363.mis1, potential  
; NAME/KEY: misc.binding  
; LOCATION: 502..521  
; OTHER INFORMATION: 12-454-363.mis2, potential complement  
; NAME/KEY: primer bind  
; LOCATION: 139..158  
; OTHER INFORMATION: upstream amplification primer  
; NAME/KEY: primer bind  
; LOCATION: 634..652  
; OTHER INFORMATION: downstream amplification primer, complement  
; NAME/KEY: misc.binding  
; LOCATION: 489..513  
; OTHER INFORMATION: 12-454-363 potential probe  
; NAME/KEY: misc.feature  
; LOCATION: 674..679,881..882,892..893  
; OTHER INFORMATION: n=a, g, c or t  
US-09-671-317-14

Query Match 1.9%; Score 42.4; DB 4; Length 929;  
Best Local Similarity 52.4%; Pred. No. 0.014;  
Matches 88; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 2115 AACCTGAGGAAGAACCCCTCGAGAGAGAAACCGCAAGCGAGAACCCAGAGCTCTCCAA 2174  
DB 896 AGCNWCACGGGANNCCATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 837  
QY 2175 AACCAACAGAGAACCCAGAGGAATCACCAGAGGAATCAGAGAACCTCAGGTGCGAGACTG 2234  
DB 836 AAG 777  
QY 2235 AAAGCTTGAGAGAAAACCTGAGAGAGCTGGAAGATTACTTGAAAAA 2282  
DB 776 AAG 729

## RESULT 35

US-09-286-981B-34  
; Sequence 34, Application US/09286981B  
; Patent No. 6503511  
; GENERAL INFORMATION:

; APPLICANT: Wizemann, Theresa M.  
; APPLICANT: Koenig, Scott  
; APPLICANT: Johnson, Leslie S  
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines  
; FILE REFERENCE: 469201-396  
; CURRENT APPLICATION NUMBER: US/09/286,981B  
; CURRENT FILING DATE: 1999-04-06  
; PRIOR APPLICATION NUMBER: US 60/085,743  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 1236  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:cdna derived  
; OTHER INFORMATION: from the genome Streptococcus pneumoniae  
US-09-286-981B-34

Query Match 1.9%; Score 42.4; DB 4; Length 1236;  
Best Local Similarity 54.5%; Pred. No. 0.017;  
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 2083 AAAACCAAGCGAGGAGAAACCTCAGACAGAGAAACCTTGAGGAGAAACCCCTCGAGAAGA 2142  
DB 1080 AAAAGCAGCAGAGAGAGATTAAGTTAAAGAAAGAAACACAGCTGAACACCAACCGCGCC 1139  
QY 2143 GAAACCGCAAGCGAGAGAAACCGAGAGTCTCCAAACCAACCGAGGAAACCGAGGAATCACC 2202  
DB 1140 GGCTCTCTCAACCAAGAAACCGAGCTCCAGCTCCAAAACCGAGAGATCCAGCTGAACACC 1199  
QY 2203 AGAGGAATCAGAGAAACCTCAGCTCGAGACTGAAAA 2238  
DB 1200 AAAAGCAGAGAAACCGAGCTGATCAACAGCTGAAGA 1235

## RESULT 36

US-09-286-981B-23  
; Sequence 23, Application US/09286981B  
; Patent No. 6503511  
; GENERAL INFORMATION:  
; APPLICANT: Wizemann, Theresa M.  
; APPLICANT: Koenig, Scott  
; APPLICANT: Johnson, Leslie S  
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines  
; FILE REFERENCE: 469201-396  
; CURRENT APPLICATION NUMBER: US/09/286,981B  
; CURRENT FILING DATE: 1999-04-06  
; PRIOR APPLICATION NUMBER: US 60/085,743  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 1338  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:cdna derived  
; OTHER INFORMATION: from the genome of Streptococcus pneumoniae  
US-09-286-981B-23

Query Match 1.9%; Score 42.4; DB 4; Length 1338;  
Best Local Similarity 54.5%; Pred. No. 0.018;  
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 2083 AAAACCAAGCGAGGAGAAACCTCAGACAGAGAAACCTTGAGGAGAAACCCCTCGAGAAGA 2142  
DB 1182 AAAAGCAGCAGAGAGAGATTAAGTTAAAGAAAGAAACACAGCTGAACACCAACCGCGCC 1241  
QY 2143 GAAACCGCAAGCGAGAGAAACCGAGAGTCTCCAAACCAACCGAGGAAACCGAGGAATCACC 2202  
DB 1242 GGCTCCAAAACAGAGAAACCGAGCTCCAGCTCCAAAACCGAGAGATCCAGCTGAACACC 1301

Qy 2203 AGAGGAATCAGAAGAACCTCAGGTCGAGACTGAAA 2238  
D<sub>b</sub> 1302 AAAAGCAGAAAAACCAGCTGATCAACAAGCTGAAGA 1337

```

1  RESULT 37
2  US-09-221-017B-27
3  ; Sequence 27, Application US/09221017B
4  ; Patent No. 6444799
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Ross, Bruce C.
7  ; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
8  ; NUMBER OF SEQUENCES: 1120
9  ; CORRESPONDENCE ADDRESS:
10 ; ADDRESSEE: MORRISON & FOERSTER
11 ; STREET: 755 PAGE MILL ROAD
12 ; CITY: Palo Alto
13 ; STATE: CA
14 ; COUNTRY: USA
15 ; ZIP: 94304-1018
16 ; COMPUTER READABLE FORM:
17 ; MEDIUM TYPE: Diskette
18 ; COMPUTER: IBM Compatible
19 ; OPERATING SYSTEM: Windows
20 ; SOFTWARE: FastSeq for Windows Version 2.0b
21 ; CURRENT APPLICATION DATA:
22 ; APPLICATION NUMBER: US/09/221,017B
23 ; FILING DATE: 23-DEC-1998
24 ; CLASSIFICATION:

```

	Query Match	1.9%	Score 42.4;	DB 4;	Length 5295;
	Best Local Similarity	54.5%;	Pred. No. 0.047;		
	Matches	85;	Conservative	71;	Indels 0; Gaps 0;
QY	2093	GAGGAGAAACCTTCAGACAGAAAAACCTTGAGAGAGAAACCCCTCGAGAGAGAGAAACCGCAA	2152		
Db	333	GAAGATTAAGCCCAAGGAAGAAAAACAAACGACAGAGAGCTTAACCGGAGGAACCCAAA	392		

QY	2153	AGCAGAAA	CCAGAGTCTC	CAAAACCA	CAGAGGAAC	CGAGAATCA	CCAGAGGAATCA	2212
Db	393	CCGAGAGCC	GTACAGCGGCT	CCCGTTC	CAAGCCGT	CGAAAA	CCTGTGGATAAA	452
QY	2213	GAGAAACCT	CAGGTCGAGAC	TGAAAAAGG	TTGAAGAA			2248
Db	453	CCACAGCAAC	CCGGTCATGACA	CAGAAAA	CCTCAAGAA			488

RESULT 38

US-09-543-681A-1562

US-09-543-681A-1562

Sequence 1562, Application US/09543681A

Patent No. 6605709

GENERAL INFORMATION:

APPLICANT: GARY BRETON

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.1002-001

CURRENT APPLICATION NUMBER: US/09/543,681A

CURRENT FILING DATE: 2000-04-05

PRIOR APPLICATION NUMBER: US 60/128,706

PRIOR FILING DATE: 1999-04-09

NUMBER OF SEQ ID NOS: 8344

SEQ ID NO 1562

LENGTH: 2766

TYPE: DNA

ORGANISM: Proteus mirabilis

US-09-543-681A-1562

Query Match	1.8%	Score 42.2;	DB 4;	Length 2766;
Best Local Similarity	47.8%;	Pred. No. 0.034;		
Matches 122;	Conservative 0;	Mismatches 133;	Indels 0;	Gaps 0;
QY	2035	CCATGTTCTCAAGAAACAA	AAATGGTCTCAGCTGATACCAATCAACGCGAA	AAACCGACGGA 2094
Db	420	CGAAGCTGAAGAAAAACG	CAAAACGTGAAGCTCAGAGCCCAAGAAAAAGCGTGAAGAAA	479
QY	2095	GGAGAAACCTTCACACAG	AAAAAACCTGAGGAAGAAACCCCTCGAGAACAGAAAAACCGCAAG	2154
Db	480	GGCAAAACGTGAAGCTGA	AGAGCAAAACCTGAAGCAGCGGAAATTAGCTTAA	CGCCGGAAGC 539
QY	2155	CGAGAAACACGAGTCTCC	AAAAACCAACAGAGGAAACAGAGAAATCA	CCAGAGGAATCAGA 2214
Db	540	AGCGGAAAAGACAAAGTG	NAACCAAAACGAGAAACCCAAAGCTCATAGGCTGATCAGGA	599
QY	2215	AGAACCTCAGGTCGAGACT	GTAAAGGTTGAAGAAAAA	CTGAGAGAGGCTCGAAGATTTACT 2274
Db	600	AAAAGCACGTCGCATAGC	CGCAACAGGCTGAAGTAAACCGTAAACAGAAAGCTCAGCG	659
QY	2275	TGGAATAATCCAGGA	2289	
Db	660	TCGTAAGGCTGAAGA	674	

```

RESULT 39
US-09-286--981B-22
; Sequence 22, Application US/09286981B
; Patent No. 6503511
; GENERAL INFORMATION:
; APPLICANT: Wizenmann, Theresa M.
; APPLICANT: Koenig, Scott
; APPLICANT: Johnson, Leslie S
; TITLE OF INVENTION: Derivatives of Choline Binding Proteins for Vaccines
; FILE REFERENCE: 463201-396
; CURRENT APPLICATION NUMBER: US/09/286, 981B
; CURRENT FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: US 60/085,743
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 1239

```

TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: cDNA derived  
from the genome of Streptococcus pneumoniae.  
US-09-286-981B-22

Query Match 1.8%; Score 42; DB 4; Length 1239;  
Best Local Similarity 50.0%; Pred. No. 0.023;  
Matches 105; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 2039 GTTCAAGAAACAAATAATGGTCAAGCTGATACCAATCAACGCGAAACCAAGCGAGGAG 2098  
Db 1030 GATCTTAAAGACGAGAGAGAGAGCTTAACGAAAGCGAGAGAGATTAAGTTAA 1089  
QY 2099 AAACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAGAGAGAAACCGCAAGCGAG 2158  
Db 1090 GAAAAACAGCTGACACACCAACACAGCGCGGCTCTCAACCGAGAAAAACCACTGAA 1149  
QY 2159 AAACAGAGTCTCCAAACCAACAGAGGAAACAGAGAAATCAACGAGGAATCAGAGAA 2218  
Db 1150 GAGCTGAGATCCAGCTCCAGCTCCAAACCAAGAGAGAGAGCTGAACCAACCAAAAGCA 1209  
QY 2219 CCTCAGGTCGAGACTGAAAGGTTGAAGAA 2248  
Db 1210 GAAAAACAGATGATCAACAGCTGAAGAA 1239

RESULT 40  
US-09-286-981B-28  
Query Match 1.8%; Score 42; DB 4; Length 1275;  
Best Local Similarity 50.0%; Pred. No. 0.023;  
Matches 105; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 2039 GTTCAAGAAACAAATAATGGTCAAGCTGATACCAATCAACGCGAAACCAAGCGAGGAG 2098  
Db 1066 GATCTTAAAGACGAGAGAGAGAGCTTAACGAAAGCGAGAGAGATTAAGTTAA 1125  
QY 2099 AAACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAGAGAGAAACCGCAAGCGAG 2158  
Db 1126 GAAAAACAGCTGACACACCAACACAGCGCGGCTCTCAACCGAGAAAAACCACTGAA 1185  
QY 2159 AAACAGAGTCTCCAAACCAACAGAGGAAACAGAGAAATCAACGAGGAATCAGAGAA 2218  
Db 1186 GAGCTGAGATCCAGCTCCGCGACCAAAACCAAGAGAGAGAGCTGAACCAACCAAAAGCA 1245  
QY 2219 CCTCAGGTCGAGACTGAAAGGTTGAAGAA 2248  
Db 1246 GAAAAACAGATGATCAACAGCTGAAGAA 1275

Query Match 1.8%; Score 42; DB 4; Length 1275;  
Best Local Similarity 50.0%; Pred. No. 0.023;  
Matches 105; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 2039 GTTCAAGAAACAAATAATGGTCAAGCTGATACCAATCAACGCGAAACCAAGCGAGGAG 2098  
Db 1066 GATCTTAAAGACGAGAGAGAGAGCTTAACGAAAGCGAGAGAGATTAAGTTAA 1125  
QY 2099 AAACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAGAGAGAAACCGCAAGCGAG 2158  
Db 1126 GAAAAACAGCTGACACACCAACACAGCGCGGCTCTCAACCGAGAAAAACCACTGAA 1185  
QY 2159 AAACAGAGTCTCCAAACCAACAGAGGAAACAGAGAAATCAACGAGGAATCAGAGAA 2218  
Db 1186 GAGCTGAGATCCAGCTCCGCGACCAAAACCAAGAGAGAGAGCTGAACCAACCAAAAGCA 1245  
QY 2219 CCTCAGGTCGAGACTGAAAGGTTGAAGAA 2248  
Db 1246 GAAAAACAGATGATCAACAGCTGAAGAA 1275

RESULT 41  
US-09-461-697-193  
Query Match 1.8%; Score 41.4; DB 3; Length 696;  
Best Local Similarity 51.5%; Pred. No. 0.023;  
Matches 121; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

QY 2046 GAAACAAAAATGCTCAAGCTGATACCAATCAACGCGAAACCAAGCGAGGAGAAACCTC 2105  
Db 392 GAAAGGAAATGAGAGAGATGGAAGAGAGAGAGAGAGATGAAAGAGAGAGAGAGACA 451  
QY 2106 AGACAGAAAACTCGAGAGAGAAACCCCTCGAGAGAGAGAAACCCCAAGCGAGAAACCGAG 2165  
Db 452 GAAAGAGAAACGAGAGTTGGAAAAAGAGAAATGAAGATGAAAGAGAGAGAGATGAAAAAG 511  
QY 2166 AGTCTCCAAAAACCAACAG 2225  
Db 512 AGGGGAAAGATGTAAGATCAAG 568  
QY 2226 TCAGAGCTGAAAGAGTTGAAG 2280  
Db 569 ATGAGGTGGAATGAG 623

Query Match 1.8%; Score 41.4; DB 3; Length 696;  
Best Local Similarity 51.5%; Pred. No. 0.023;  
Matches 121; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

QY 2046 GAAACAAAAATGCTCAAGCTGATACCAATCAACGCGAAACCAAGCGAGGAGAAACCTC 2105  
Db 392 GAAAGGAAATGAGAGAGATGGAAGAGAGAGAGAGAGATGAAAGAGAGAGAGAGACA 451  
QY 2106 AGACAGAAAACTCGAGAGAGAAACCCCTCGAGAGAGAGAGAAACCCCAAGCGAGAAACCGAG 2165  
Db 452 GAAAGAGAAACGAGAGTTGGAAAAAGAGAAATGAAGATGAAAGAGAGAGAGATGAAAAAG 511  
QY 2166 AGTCTCCAAAAACCAACAG 2225  
Db 512 AGGGGAAAGATGTAAGATCAAG 568  
QY 2226 TCAGAGCTGAAAGAGTTGAAG 2280  
Db 569 ATGAGGTGGAATGAG 623

RESULT 42  
US-09-461-697-191  
Query Match 1.8%; Score 41.4; DB 3; Length 696;  
Best Local Similarity 51.5%; Pred. No. 0.023;  
Matches 121; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

QY 2046 GAAACAAAAATGCTCAAGCTGATACCAATCAACGCGAAACCAAGCGAGGAGAAACCTC 2105  
Db 392 GAAAGGAAATGAGAGAGATGGAAGAGAGAGAGAGAGATGAAAGAGAGAGAGAGACA 451  
QY 2106 AGACAGAAAACTCGAGAGAGAAACCCCTCGAGAGAGAGAGAAACCCCAAGCGAGAAACCGAG 2165  
Db 452 GAAAGAGAAACGAGAGTTGGAAAAAGAGAAATGAAGATGAAAGAGAGAGAGATGAAAAAG 511  
QY 2166 AGTCTCCAAAAACCAACAG 2225  
Db 512 AGGGGAAAGATGTAAGATCAAG 568  
QY 2226 TCAGAGCTGAAAGAGTTGAAG 2280  
Db 569 ATGAGGTGGAATGAG 623

Query Match 1.8%; Score 41.4; DB 3; Length 696;  
Best Local Similarity 51.5%; Pred. No. 0.023;  
Matches 121; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

QY 2046 GAAACAAAAATGCTCAAGCTGATACCAATCAACGCGAAACCAAGCGAGGAGAAACCTC 2105  
Db 392 GAAAGGAAATGAGAGAGATGGAAGAGAGAGAGAGAGATGAAAGAGAGAGAGAGACA 451  
QY 2106 AGACAGAAAACTCGAGAGAGAAACCCCTCGAGAGAGAGAGAAACCCCAAGCGAGAAACCGAG 2165  
Db 452 GAAAGAGAAACGAGAGTTGGAAAAAGAGAAATGAAGATGAAAGAGAGAGAGATGAAAAAG 511  
QY 2166 AGTCTCCAAAAACCAACAG 2225  
Db 512 AGGGGAAAGATGTAAGATCAAG 568  
QY 2226 TCAGAGCTGAAAGAGTTGAAG 2280  
Db 569 ATGAGGTGGAATGAG 623









## ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 134:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 12665 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-961-527-134

```
Query Match      1.8%; Score 41.4; DB 4; Length 12665;
Best Local Similarity 51.3%; Pred.No. 0.17;
Matches 96; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 2083 AAAACCAAGCGAGGAGAAACCTCAGACAGAAAAAACCCTGAGGAGAAACCCCTCGAGAAGA 2142
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2143 GAAACCGCAAGCGAGAAACCGAGAGTCTCCAAAACCAACAGAGGAAACCGAGAATCAAC 2202
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1923 GGCTCCAAAAGCAGAAAAACCGAGCTCCAGTCCAAAACCGAGAGATCCAGCTGAACAACC 1982
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2203 AGAGGAATCAGAGAACCTCAGGTCGAGACTGAAAAGGTTGAAGAAAAACTGAGAGAGGC 2262
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1983 AAAAGCAGAAAAACCGAGCTGATCAACAGCTGAAGAAGACTATGCTCTGTAGATCAGAAGA 2042
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2263 TGAGAT 2269
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2043 AGATAT 2049
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: May 1, 2004, 04:16:53  
Job time : 175 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 1, 2004, 06:21:35 ; Search time 776 Seconds  
(without alignments)  
4447.684 Million cell updates/sec

Title: US-09-765-272A-66

Perfect score: 4019

Sequence: 1 CSYEGLRHQAGQVKKESNRV.....TEKVEEKLREADLLGKIQD 763

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delopt 6.0 , Delopt 7.0

Searched: 2936184 seqs, 2261732022 residues

Total number of hits satisfying chosen parameters: 5872368

Minimum DB seq length: 0

Maximum Match 100%

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09765272/runat\_30042004\_081052\_11633/app\_query.fasta\_1.903  
-DB=Published Applications NA -QFMT=fastap -SURFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bats -START=1 -END=1 -MATRIX=blomsum62  
-TRANS=human40.cdi -LIST=150 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-MAXLEN=2000000000 -USER=US09765272@cgn\_1\_232@runat\_30042004\_081052\_11633  
-NCFU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:  
9: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/prodata/1/pubpna/US09A\_NEW\_PUB.seq:  
14: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:  
17: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:  
18: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:  
19: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99.9	4.013	2290	9	US-09-765-272-65
2	3968.5	98.7	2481	10	US-09-769-787-206
3	3919	97.5	2531	13	US-10-412-862-11
4	3919	97.5	2531	16	US-10-412-850-11
5	3919	97.5	2531	16	US-10-387-783-11
6	3665	91.2	2639	10	US-09-884-465A-5
7	3644.5	90.7	2531	13	US-10-412-862-5
8	3644.5	90.7	2531	16	US-10-412-850-5
9	3644.5	90.7	2531	16	US-10-387-783-5
10	3641.5	90.6	8195	13	US-10-158-844-94
11	3588.5	89.3	2523	10	US-09-884-465A-3
12	3588.5	89.3	2647	10	US-09-884-465A-4
13	2658.5	66.1	2451	13	US-10-412-862-9
14	2658.5	66.1	2451	16	US-10-412-850-9
15	2658.5	66.1	2451	16	US-10-387-783-9
16	2649.5	65.9	2389	9	US-09-765-272-55
17	2424	60.3	2359	13	US-10-158-844-243
18	1377	34.3	973	13	US-10-158-844-355
19	1228.5	30.6	3120	10	US-09-884-465A-1
20	1228.5	30.6	5048	10	US-09-884-465A-2
21	1202.5	29.9	6867	13	US-10-158-844-192
22	1200	29.9	1684	13	US-10-158-844-258
23	1167.5	29.0	1455	10	US-09-769-787-246
24	1167.5	29.0	1455	10	US-09-769-744A-23
25	1167.5	29.0	1455	13	US-10-412-862-7
26	1167.5	29.0	1455	16	US-10-412-850-7
27	1167.5	29.0	1455	16	US-10-387-783-7
28	1141.5	28.4	1342	9	US-09-765-272-181
29	1023.5	25.5	5215	10	US-09-252-088-13
30	1023.5	25.4	5215	16	US-10-340-792-13
31	1019.5	25.4	2469	10	US-09-769-736-17
32	980	24.4	2528	10	US-09-884-465A-9
33	877	21.8	1146	10	US-09-769-736-23
34	827.5	20.6	819	10	US-09-884-465A-257
35	827.5	20.6	819	16	US-10-324-143-10
36	827.5	20.6	2721	16	US-10-324-143-6
37	822.5	20.5	816	16	US-10-324-143-4
38	822.5	20.5	816	16	US-10-324-143-5
39	542.5	13.5	1455	10	US-09-769-736-71
40	512	12.7	841	9	US-09-452-599-34
41	512	12.7	841	15	US-10-121-120-34
42	190.5	4.7	3466	16	US-10-104-047-536
43	186.5	4.6	3494	13	US-10-087-192-233
44	186.5	4.6	5316	16	US-10-159-563-295
45	184.5	4.6	3611	9	US-09-974-298-113
46	180.5	4.5	5126	15	US-10-232-561-9
47	179.5	4.5	3570	16	US-10-369-493-45846
48	178	4.4	2522	9	US-09-842-553-98
49	177.5	4.4	8588	9	US-09-816-669A-13
50	177.5	4.4	8588	13	US-10-342-887-1335
51	177.5	4.4	8588	13	US-10-172-118-1335
52	175	4.4	3891	13	US-10-282-122A-35508
53	169.5	4.2	2232	15	US-10-087-464-45
54	169	4.2	3825	13	US-10-283-122A-7653
55	168.5	4.2	3284	15	US-10-128-714-7305
56	168.5	4.2	5142	13	US-10-087-192-230
57	168	4.2	2883	15	US-10-128-714-7193
58	168	4.2	2939	15	US-10-128-714-6193
59	168	4.2	4225	15	US-10-128-714-193
60	168	4.2	4939	15	US-10-128-714-5193
61	167.5	4.2	3336	9	US-09-815-242-8557
62	167	4.2	4088	10	US-09-814-353-20949
63	165	4.1	2392	9	US-09-216-393-331
64	165	4.1	2392	15	US-10-321-856-331
65	165	4.1	2393	9	US-09-216-393-329
66	165	4.1	2393	15	US-10-321-856-329
67	165	4.1	3379	15	US-10-128-714-6305
68	165	4.1	5379	15	US-10-128-714-5305
69	164.5	4.1	9072	16	US-10-396-122-87
70	164	4.1	2865	13	US-10-424-599-24135
71	164	4.1	3003	15	US-10-128-714-2305
72	164	4.1	3809	14	US-10-001-870-66

```

73 163.5 4.1 2561 9 US-09-880-107-2435 Sequence 2435, Ap
74 163.5 4.1 2561 13 US-10-342-887-777 Sequence 777, App
75 163.5 4.1 2561 13 US-10-172-118-777 Sequence 777, App
76 163.5 4.1 2561 16 US-10-170-385-474 Sequence 474, App
77 163.5 4.1 2561 16 US-10-369-493-29245 Sequence 29245, A
78 163.5 4.1 4200 13 US-09-797-385-1 Sequence 1, Appli
79 162.5 4.0 5039 15 US-10-128-714-305 Sequence 305, App
80 162.5 4.0 3039 15 US-10-128-714-1305 Sequence 1305, Ap
81 161.5 4.0 2479 16 US-10-108-260A-1966 Sequence 1966, Ap
82 160.5 4.0 4522 8 US-08-781-986A-39 Sequence 39, Appl
83 160.5 4.0 4522 13 US-10-329-624-39 Sequence 39, Appl
84 159.5 4.0 1766 13 US-09-298-523B-74 Sequence 74, Appl
85 159.5 4.0 1766 13 US-09-748-875-74 Sequence 74, Appl
86 159.5 4.0 80161 17 US-10-329-148A-1 Sequence 1, Appli
87 159.5 4.0 3238 16 US-10-159-563-209 Sequence 209, App
88 159.5 4.0 81940 9 US-09-759-508B-1 Sequence 1, Appli
89 159.5 4.0 81940 10 US-09-960-706-1092 Sequence 1092, Ap
90 159.5 4.0 81940 10 US-08-873-319-724 Sequence 724, App
91 158.5 3.9 2415 15 US-10-156-761-4284 Sequence 4284, Ap
92 158.5 3.9 3953 15 US-10-232-561-11 Sequence 11, Appl
93 158.5 3.9 9025608 15 US-10-156-761-1 Sequence 1, Appli
94 158.5 3.9 1278 15 US-10-254-995-29 Sequence 29, Appl
95 158.5 3.9 3290 15 US-10-084-817-286 Sequence 286, App
96 157.5 3.9 1272 15 US-10-254-995-31 Sequence 31, Appli
97 157.5 3.9 3132 9 US-09-946-805-3 Sequence 3, Appli
98 157.5 3.9 18717 15 US-10-156-761-927 Sequence 927, App
99 156.5 3.9 1276 15 US-10-254-995-30 Sequence 30, Appl
100 156.5 3.9 2082 10 US-09-769-787-197 Sequence 197, App
101 156.5 3.9 12665 13 US-10-158-844-134 Sequence 134, App
102 154.5 3.8 2655 13 US-10-282-122A-10255 Sequence 10255, A
103 154.5 3.8 28473 13 US-10-158-844-83 Sequence 83, Appl
104 153 3.8 4533 16 US-10-418-861B-54 Sequence 54, Appl
105 153 3.8 4851 13 US-10-282-122A-27409 Sequence 27409, A
106 153 3.8 9399 16 US-10-396-122-93 Sequence 93, Appl
107 153 3.8 580073 15 US-10-205-220-1 Sequence 1, Appli
108 152.5 3.8 2169 15 US-10-128-714-2193 Sequence 2193, Ap
109 152.5 3.8 2225 15 US-10-128-714-1193 Sequence 1193, Ap
110 152.5 3.8 2769 16 US-10-108-260A-280 Sequence 280, App
111 152 3.8 3990 10 US-09-769-744A-165 Sequence 165, App
112 151.5 3.8 7095 13 US-10-205-331-67 Sequence 67, Appl
113 151.5 3.8 10656 13 US-10-263-329-42 Sequence 42, Appli
114 151 3.8 3384 13 US-09-797-385-5 Sequence 5, Appli
115 151 3.8 10711 13 US-10-158-844-145 Sequence 145, App
116 150.5 3.7 1338 15 US-10-254-995-26 Sequence 26, Appl
117 150.5 3.7 2405 13 US-09-298-523B-70 Sequence 70, Appl
118 150.5 3.7 2405 13 US-09-748-875-70 Sequence 70, Appl
119 149 3.7 1325 10 US-08-814-353-19142 Sequence 19142, A
120 149 3.7 2550 15 US-10-032-585-6026 Sequence 6026, Ap
121 149 3.7 3294 13 US-09-797-385-7 Sequence 7, Appli
122 149 3.7 3444 16 US-10-369-493-40475 Sequence 40475, A
123 149 3.7 3492 13 US-09-797-385-9 Sequence 9, Appli
124 149 3.7 3847 13 US-10-320-797-237 Sequence 237, App
125 148.5 3.7 2122 13 US-10-424-599-119619 Sequence 119619,
126 148 3.7 4512 16 US-10-369-493-27634 Sequence 27634, A
127 148 3.7 6138 13 US-10-282-122A-38279 Sequence 38279, A
128 147.5 3.7 9071 16 US-10-115-831-127 Sequence 127, App
129 147.5 3.7 9212 16 US-10-115-831-126 Sequence 126, App
130 147 3.7 1855 9 US-09-864-761-19708 Sequence 19708, A
131 147 3.7 1872 15 US-10-029-386-20427 Sequence 20427, A
132 147 3.7 1958 9 US-09-864-761-2927 Sequence 2927, Ap
133 147 3.7 5295 13 US-10-194-163-37 Sequence 37, Appl
134 146.5 3.6 1969 10 US-09-056-019-13 Sequence 13, Appl
135 146.5 3.6 1989 16 US-10-104-047-745 Sequence 745, App
136 146.5 3.6 3851 16 US-10-369-493-26948 Sequence 26948, A
137 146.5 3.6 5346 9 US-09-738-877-2 Sequence 2, Appli
138 146.5 3.6 6608 9 US-09-738-877-1 Sequence 1, Appli
139 146.5 3.6 6608 9 US-09-880-107-3439 Sequence 3439, Ap
140 146.5 3.6 6608 10 US-09-873-367C-275 Sequence 275, App
141 146.5 3.6 6608 13 US-10-211-462-43 Sequence 43, Appl
142 146.5 3.6 6608 16 US-10-428-487-15 Sequence 15, Appl
143 146.5 3.6 10567 16 US-10-144-194A-61 Sequence 61, Appl
144 146.5 3.6 9025608 15 US-10-156-761-1 Sequence 1, Appli
145 146 3.6 2152 10 US-09-764-891-7678 Sequence 7678, Ap

```

Sequence 2435, Ap  
Sequence 777, App  
Sequence 777, App  
Sequence 474, App  
Sequence 29245, A  
Sequence 1, Appli  
Sequence 305, App  
Sequence 1305, Ap  
Sequence 1966, Ap  
Sequence 39, Appl  
Sequence 39, Appl  
Sequence 74, Appl  
Sequence 74, Appl  
Sequence 1, Appli  
Sequence 209, App  
Sequence 1, Appli  
Sequence 1092, Ap  
Sequence 724, App  
Sequence 4284, Ap  
Sequence 11, Appl  
Sequence 1, Appli  
Sequence 29, Appl  
Sequence 286, App  
Sequence 31, Appli  
Sequence 3, Appli  
Sequence 927, App  
Sequence 30, Appl  
Sequence 197, App  
Sequence 134, App  
Sequence 10255, A  
Sequence 83, Appl  
Sequence 54, Appl  
Sequence 27409, A  
Sequence 93, Appl  
Sequence 1, Appli  
Sequence 2193, Ap  
Sequence 1193, Ap  
Sequence 280, App  
Sequence 165, App  
Sequence 67, Appl  
Sequence 42, Appli  
Sequence 5, Appli  
Sequence 145, App  
Sequence 26, Appl  
Sequence 70, Appl  
Sequence 70, Appl  
Sequence 19142, A  
Sequence 6026, Ap  
Sequence 7, Appli  
Sequence 40475, A  
Sequence 9, Appli  
Sequence 237, App  
Sequence 119619,  
Sequence 27634, A  
Sequence 38279, A  
Sequence 127, App  
Sequence 126, App  
Sequence 19708, A  
Sequence 20427, A  
Sequence 2927, Ap  
Sequence 37, Appl  
Sequence 13, Appl  
Sequence 745, App  
Sequence 26948, A  
Sequence 2, Appli  
Sequence 1, Appli  
Sequence 3439, Ap  
Sequence 275, App  
Sequence 43, Appl  
Sequence 15, Appl  
Sequence 61, Appl  
Sequence 1, Appli  
Sequence 7678, Ap

Sequence 38, Appl  
Sequence 20951, A  
Sequence 6631, Ap  
Sequence 1, Appli  
Sequence 14, Appl

146 146 3.6 7058 15 US-10-287-218-38  
147 145 3.6 2334 13 US-10-282-122A-20951  
148 145 3.6 2337 9 US-09-815-242-6631  
149 145 3.6 5065 15 US-10-339-782-1  
150 145 3.6 7554 16 US-10-414-692-14

ALIGNMENTS

RESULT 1  
US-09-765-272-65  
; Sequence 65, Application US/09765272  
; Patent No. US20020061545A1  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; FILING DATE: 22-Jan-2001  
; APPLICATION NUMBER: US/09/765,272  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 35,373  
; REFERENCE/DOCKET NUMBER: FB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 65:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2290 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-09-765-272-65

Alignment Scores:  
Pred. No.: 0 Length: 2290  
Score: 4013.00 Matches: 763  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 99.85% Indels: 0  
DB: 9 Gaps: 0

US-09-765-272A-66 (1-763) x US-09-765-272-65 (1-2290)

QY 1 CysSerTyrluLeuGlyArgHisGlnAlaGlyGlnVallylsGluSerAsnArgVal 20  
DB 2 TGTTCCTATGAATGGTGGTCGCCACCAAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGTT 61  
QY 21 SerTyrlleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40  
DB 62 TTTTATATAGATGGTGATCAGGCTGGTCAAAAGGCAGAAACTTGACACCAGATGAAGTC 121  
QY 41 SerLysArgGluGlyIleAsnAlaGluGln\*\*\*ValIleLysIleThrAspGlnGlyTyr 60  
DB 122 ACTAAGAGGGGGGGGATCAACGCCGACAAATGNTATCAAGATTACGGATCAAGTTAT 181

61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaile 80  
 182 GTGACCTCTCATGGAGACCATATCATTAATAATGGCAAGTTCCTTATGATGCCATC 241  
 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100  
 242 ATCAGTGAAGAGCTCCTCATGAAAGATCCGAATATATCAGTTGAAGGATTTCAGACATTGTC 301  
 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120  
 302 AATGAATCAAGGGTGTATGTCTATTAAGTAAACGGTAATACTATAGTNTACCTTAAG 361  
 121 AspAlaIleHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140  
 362 GATGCAGCTCATGGGATAATATTCGGACAAAAGAGAGATTAAACGTCAAGAGCAGGAA 421  
 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160  
 422 CGCAGTCAATAATCATAACTCAAGAGCAGATAATATGCTGTGCTCAGCCAGAGCCCAAGGA 481  
 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180  
 482 CGTTATACACGAGATGATGGGTATATCTTCATATGCATCTCATATCATTTGAGGACACGGGT 541  
 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200  
 542 GATGCTTATATCTCTCAGCGCGACCATTAACATTACATTCCTTAAGATGAGTTATCA 601  
 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLysGlnGlySerArgProSer 220  
 602 GCTAGCGAGTTAGTCTCAGAGAGCTATATGGAATGGGAAGCAGGAGATCTCGTCTCTCT 661  
 221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240  
 662 TCAAGTCTTAGTTAATGCAATATCCAGCTCAACCAAGATTGTCAGAGAACCAACATCTG 721  
 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260  
 722 ACTGTCACTCCAACTTATCATCAAAATCAAGGGAACAACTTCAAGCCCTTTTACGTGAA 781  
 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280  
 782 TTGTATGCTAAACCCCTTATCAGACGCCATGTGGAATCTGATGGCTTATTTTCGACCCA 841  
 281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300  
 842 CGCAGAAATCACAAAGTCGAACCCGCGAGAGTGTAGCTGTCCCTCATGGTAACCAATTACCAC 901  
 301 PheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeu 320  
 902 TTTATCCCTTATGAACAATGTCTGNAATGGAAAAACGAATGTCTGTATATATCCCTTT 961  
 321 ArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSer 340  
 962 CGTTATCGTTCAAACCATTTGGGTACCAAGTTCAGAGTTCAGAACCAACCAAGTCCACAAATCG 1021  
 341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360  
 1022 ACTCCGGAACTTAGTCCAAAGTCCGAACCTCGCAACATCTCCAAATCTCTCAACGAGCTCCAAAGCAAT 1081  
 361 ProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380  
 1082 CCNATGTAGAAATTTGTCTAAAGAGTGTTCGAAAGATGAGCGCATGTTATGTTCTT 1141  
 381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400  
 1142 GAGGAGATGGAGTTCTCGTTATATCCAGCCCAAGATCTTTCACGAGAAACAGCAGCA 1201  
 401 GlyIleAspSerLysLeuAlaLysGlnGlnSerHisGlyLysLeuGlyAlaLysLys 420  
 1202 GGCATTGATAGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAAAGCTAGCAGCTAAGAAA 1261

421 ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440  
 1262 ACTGACCTCCCATCTAGTATGATCGAGAAATTTTACATAATAGGCTTATGACTTACTAGCAAGA 1321  
 441 IleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460  
 1322 ATTCACCAAGATTTACTTTGATAATAAGGTCACAAAGTTGATTTTGGAGCTTTGGATAAC 1381  
 461 LeuLeuGluArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeu 480  
 1382 CTGTTGGACACACTCAAGATGTCNCAAGTGATAAGTCAAGTTAGTGGANGATATCTT 1441  
 481 AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr 500  
 1442 GCCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTAGGAAAAACCAATTCGCAATATACC 1501  
 501 TyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGly 520  
 1502 TACACTGATGATGAGATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACAAACAGAACGCGT 1561  
 521 TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis 540  
 1562 TATATCTTTGATCTCTCGTGATATACCAAGTGATGAGGGGATGCCATGTAACCTCCACAT 1621  
 541 MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAla 560  
 1622 ATGACCCATAGCCACTGGATTAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGAGCC 1681  
 561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580  
 1682 CAGGCTTATGCTAAAGAGAAAGTTTGACCCCTCTCTCGACAGACCATCAGGATTTCAGGA 1741  
 581 AsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysVal 600  
 1742 AATACTCAGGCAAAAGGAGCAGAGAGCTATCTACACCCGCGTGAAGAGCTAAGAGGTG 1801  
 601 ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu 620  
 1802 CCATCTGATCGTATGCTTCAATCTTCAATATCTGTAGAGTCAAAAACCGTAGTTT 1861  
 621 IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeu 640  
 1862 ATCATACCTCATATGACCATTAACATAACATCAAAATTTGAGTGTGTTTACGAGGSCCTT 1921  
 641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrVal 660  
 1922 TATGAGGCACCTAAGGGGTATACTCTTTGAGGATCTTTTGGCGACTGTCAAGTACTATGTC 1981  
 661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680  
 1982 GACATCCAAACGAAACGTCGCAATTCAGATTAATGTTTTGTTAGCTAGCGACCATGTT 2041  
 681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLys 700  
 2042 CAAAGAAACAAAATGCTCAAGCTGATACCAATCAACCGGAAAAACCAAGCGAGAGAAA 2101  
 701 ProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLys 720  
 2102 CCTCAGACAGAAAAACCTGAGGAGAAACCCCTCGAGAAAGAGAAACCGCAAGCAGAGAA 2161  
 721 ProGlnSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro 740  
 2162 CCAGAGTCTCCAAAAACCAACAGAGAAACCAAGAAATCACCAGAGGAATCAGAGAAACCT 2221  
 741 GlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLys 760  
 2222 CAGTCTGAGACTGAAAGAGTTGAAGAAACTGAGAGGGCTGAAGATTACTTACTTGAAAA 2281  
 761 IleGlnAsp 763  
 2282 ATCCAGGAT 2290

RESULT 2

US-09-769-787-206  
 ; Sequence 206, Application US/09769787  
 ; Publication No. US20030091577A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Microbial Technics Limited  
 ; APPLICANT: Gilbert, Christophe FG  
 ; APPLICANT: Hansbro, Philip M  
 ; TITLE OF INVENTION: Proteins  
 ; FILE REFERENCE: PWC/P21129W0  
 ; CURRENT FILING DATE: 2001-01-26  
 ; PRIOR APPLICATION NUMBER: GB 9816337.1  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: US 60/125164  
 ; PRIOR FILING DATE: 1999-03-19  
 ; NUMBER OF SEQ ID NOS: 388  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 206  
 ; LENGTH: 2481  
 ; TYPE: DNA  
 ; ORGANISM: Streptococcus pneumoniae  
 US-09-769-787-206

Alignment Scores:  
 Pred. No.: 0 Length: 2481  
 Score: 3968.50 Matches: 754  
 Percent Similarity: 99.08% Conservative: 3  
 Best Local Similarity: 98.69% Mismatches: 6  
 Query Match: 98.74% Indels: 1  
 DB: 10 Gaps: 1

US-09-765-272A-66 (1-763) x US-09-769-787-206 (1-2481)

QY	1	CysSerTyrGluLeuGlyValArgHisGlnAlaGlyGlnValLysGluSerAsnArgVal	20
DB	58	TGTTCTATAGCTGGTGGTCCACCAAGCTGGTCAGGATAAGAAAGAGTCTAATCGAGTT	117
QY	21	SerTyrIleAspGlyValAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal	40
DB	118	GCTTATATAGATGGTGGTGGTCCAAAGGCGAAGAACTTGACACCAAGTGAAGTC	177
QY	41	SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr	60
DB	178	AGTAGAGGGGGGATCAACCGCAACAAATCGTTCATCAAGATTACGGATCAAGGTTAT	237
QY	61	ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle	80
DB	238	GTGACCTCTCATGGAGACCAATATATCTACTATAATGGCAAGTCCCTTATGATGCCATC	297
QY	81	IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLysAspSerAspIleVal	100
DB	298	ATCAGTGAAGAGCTCTCATGAAGATCCGANTTATCAGTTGAAGANTTCAGCATTTGTC	357
QY	101	AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys	120
DB	358	AATGAAATCAAGGGTGGTATGTCATCAAGGTAGACGGAATACTATGTTTACCTTAAG	417
QY	121	AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu	140
DB	418	GATGAGCTCATGCGGTAATATTCGGAACAAAGAGAGATTAAACGTCAGAGCAGGAA	477
QY	141	ArgSerHisAsnHisAsnSerArgAla---AspAsnAlaValAlaAlaArgAlaGln	159
DB	478	CGCAGTCATATCAACGGGTGAGAGTAAAGCATGATCATGCTAGTGTGACCCAGAGCCCAA	537
QY	160	GlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThr	179
DB	538	GGACGCTATACAAAGGATGATGGGTATATCTTCAATGCACTCTGATATCATTCAGGACAG	597
QY	180	GlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeu	199
DB	598	GGTGATGCTTATCGTTCCTCACGCGCACTTACCATTTACATTCCTAAGAAATGAGTTA	657

QY	200	SerAlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgPro	219
DB	658	TCAGCTAGCGAGTTAGCTGCTGCAGAAAGCCTATTGGAAATGGAAAGCAGGATCTCGTCT	717
QY	220	SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsn	239
DB	718	TCITCAAGTTCCTAGTTAATAATCAATCCAGCTCAACCAAGATTGTGAGAAACCAACAT	777
QY	240	LeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArg	259
DB	778	CTGACTGTCACTCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCTTTTACGT	837
QY	260	GluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAsp	279
DB	838	GAATTTGATGCTAAACCCCTTATCAGAACGCCATCTGGAATCTGATGGCTTATTTTCGAC	897
QY	280	ProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyr	299
DB	898	CCAGCGCAATCAAGTCCGACCGCCAGAGGTGTAGCTGTCCCTCATGGTAAACCATTAC	957
QY	300	HisPheIleProTyrGluGlnMetSerGluLeuGlyLysArgGlnAlaArgIleIlePro	319
DB	958	CACCTTATCCCTTATGAACAAATGTCTGAATGGAAAAACGAATTTGCTGATTTATTCCT	1017
QY	320	LeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerProGln	339
DB	1018	CTTTCGTTATCGTTCAAAACCATTTGGGTACCAAGATTCAAGACCAGAACCAACCAACAA	1077
QY	340	SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer	359
DB	1078	TCGACTCCGGAACTTAGTCCCAAGTCCGCAACCTCCACCAATCTCCACAGCTCCAAAGC	1137
QY	360	AsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrVal	379
DB	1138	ATCCCAATGATGAGAAATTTGGTCAAAAGAGCTGTTGAAAAGTAGGCGATGTTATGTC	1197
QY	380	PheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAla	399
DB	1198	TTTGAGGAGAAATGGAGTTTCTCGTTATATCCACCAAGGATCTTTCAGCAGAAACAGCA	1257
QY	400	AlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys	419
DB	1258	GCAGGCACTTATAGCAAACTGGCCCAAGCAGCAAGAAAGTTTATCTCATTAAGCTAGGAGTAA	1317
QY	420	LysThrAspLeuProSerSerAspAspGluPheTyrAsnLysAlaTyrAspLeuAla	439
DB	1318	AAACTGACCTCCCATCTAGTATCGAGAAATTTTACAAATAGGCTTATGACTTACTAGCA	1377
QY	440	ArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAsp	459
DB	1378	AGAATTCACCAAGATTTACTTTGATAATAAGGTCGACCAAGTTGATTTTGAAGCTTTGGAT	1437
QY	460	AsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIle	479
DB	1438	AACCTGTTGGAAGCACTCAAGGATGTCCCAAGTGATAAAGTCAAGTTAGTGTGATGATATT	1497
QY	480	LeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIle	499
DB	1498	CTTGCTCTTCTAGCTCCGATTCGTCATCCAGAAAGTTTAGGAAACCAAAATTCGCAAAAT	1557
QY	500	ThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAsp	519
DB	1558	ACCTACACTGATGATGAGATTCAGTAGCCAGTTGCGCAGGCAAGTACACACAGAGAGAC	1617
QY	520	GlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrPro	539
DB	1618	GGTTATATCTTTGATCCTCGTATATAACCAAGTATGAGGGGATGCTATGTAACTCCA	1677
QY	540	HisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAla	559
DB	1678	CATATGACCATGACCCACTGGATTAATAAGATAGTTTGTCTGAAGCTGAGAGAGCGCA	1737
QY	560	AlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSer	579

1738 GCCCAGGCTTATGCTAAAGAGAAAGGTTTGACCCCTCCTTCGACAGACCATCAGGATTCA 1797  
580 GlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLys 599  
1798 GGAATACTAGGCAAAAGAGAGAGAGATCTTACAAACCGCTGGAAGCAGCTAAGAAG 1857  
600 ValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySer 619  
1858 GTGCCACTTGATCGTATGCTTACAACTTCAATATCTAGAGGTCAAAACGGGTAGT 1917  
620 LeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGluGly 639  
1918 TTAATCATCATCTCATATGACCATTAACCATTAACATCAAAATTTGAGTGGTTTGACGAAGGC 1977  
640 LeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyr 659  
1978 CTTTATGAGGCACCTAAGGGGTATATCTTTGAGGATCTTTTGGCGACTGCAAGTACTAT 2037  
660 ValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHis 679  
2038 GTCGAACATCCAAACGAAAGCTCGCATTCAGATAATGGTTTGGTAACGCTAGCGACCAT 2097  
680 ValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGlu 699  
2098 GTTCAAAGAAACAAATAATGCTCAAGCTGATACCAATCAAAACGGGAAAAACCAAGCGAGAG 2157  
700 LysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGlu 719  
2158 AAACCTCAGACAGAAACAACTTGAGGAAGAAACCCCTCGAAGAGAAACCGCAAGCGAG 2217  
720 LysProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGlu 739  
2218 AAACAGAGTCTCCAAACCAACAGAGAACCAAGAGAAATCAACAGAGGAATCAAGAGAA 2277  
740 ProGlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuLeuGly 759  
2278 CCTCAGGTCGAGACTGAAAGGTTGAAGAAAACTGAGAGAGGCTGAAGATTACTTTGGA 2337  
760 LysIleGlnAsp 763  
2338 AAAATCCAGGAT 2349

RESULT 3  
JS-10-412-862-11  
: Sequence 11, Application US/10412862  
: Publication No. US20040052781A1  
: GENERAL INFORMATION:  
: APPLICANT: Johnson, Leslie S.  
: APPLICANT: Adamou, John E.  
: TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
: TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
: TITLE OF INVENTION: Motifs  
: FILE REFERENCE: 469201-685  
: CURRENT APPLICATION NUMBER: US/10/412,862  
: CURRENT FILING DATE: 2003-04-14  
: PRIOR APPLICATION NUMBER: 09/468,656  
: PRIOR FILING DATE: 1999-12-21  
: PRIOR APPLICATION NUMBER: 60/113,048  
: PRIOR FILING DATE: 1998-12-21  
: NUMBER OF SEQ ID NOS: 14  
: SOFTWARE: PatentIn Ver. 3.0  
: SEQ ID NO 11  
: LENGTH: 2531  
: TYPE: DNA  
: ORGANISM: Streptococcus pneumoniae  
JS-10-412-862-11  
Alignment Scores:  
Pred. No.: 0 Length: 2531  
Score: 3919.00 Matches: 745  
Percent Similarity: 98.03% Conservative: 3  
Best Local Similarity: 97.64% Mismatches: 9

Query Match: 97.51% Indels: 6  
DB: 13 Gaps: 1  
US-09-765-272A-66 (1-763) x US-10-412-862-11 (1-2531)  
Qy 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20  
Db TGTTCCTATGAGCTGGACGTTTACCAAGCTGGTCAGGATAAGAAAGAGTCTAATCGAGTT 117  
Qy 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40  
Db GCTTATATAGATGGTGAATCAGCGCTGGTCAAAAGGAGGAAACTTCACACCATGAAGTC 177  
Qy 41 SerLysArgGluGlyIleAsnAlaGluGln\*\*\*ValIleLysIleThrAspGlnGlyTyr 60  
Db AGTAAGAGGAGGGGATCAACGCCGACAAATTTGTTATCAAGATTACGGATTCAAGGTTAT 237  
Qy 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80  
Db GTGACCTCTCATGGAGACCATATCATTTACTATATATGSCAAGGTTCTTTATGATGCCATC 297  
Qy 81 IleSerGluGluLeuLeuMetLysAspProSerTyrGlnLeuLysAspSerAspIleVal 100  
Db ATCAGTCAAGAGCTCCTCATGAAAGATCCGATTTATCAGTTCAAGGATTCAGACATTGTC 357  
Qy 101 AsnGluIleLysGlyLysValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120  
Db AATGAAATCAAGGGTGGTTATGTCTAATGAGTAAACGGTAAATACTATTTTACCTTAAG 417  
Qy 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140  
Db GATGCGCTCATGCGGATAATATTCGGACAAAGAGAGATTAAACGTCAGAGCAGGAA 477  
Qy 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160  
Db CGCAGTCATATACTAACTCAAGAGCAGATATGCTGTGTGTCGACGAGAGCCCAAGGA 537  
Qy 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180  
Db CGTTATACACGATGATGGGTATATCTTCAATCATCTGATATCATGAGGACACGGGT 597  
Qy 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200  
Db GATGCTTATATCGTTCTTCACGGCGACCATACCATTTACATTCTTAAGAAATGAGTTATCA 657  
Qy 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSer 220  
Db GCTAGCGAGTTAGCTGTGCGAAGCCCTATTGSAATGGGAAGCAGGGATCTCGTCTCTCT 717  
Qy 221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240  
Db TCAAGTTCTAGTTATATGCAATCCAGTCAACCAAGATTCAGCAAGATTCAGAGAACCAATCTG 777  
Qy 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260  
Db ACTGTCACTCCACTTATCATCAAAATCAAGGGGAAACATTTTCAAGCCTTTTACGTGAA 837  
Qy 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280  
Db TTGTATGCTAAACCCCTTATCAGAACCCCATGTTGGAATCTGATGGCCTTATTTTCGACCCA 897  
Qy 281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300  
Db GCGCAAAATCAAGTCGAACCGCCAGAGGTGTAGCTGTCCCTCATGTGGTAAACCATTACCAC 957  
Qy 301 PheIleProTyrGluGlnMetSerGluLeuGlnLysArgIleAlaArgIleIleProLeu 320  
Db TTTATCCCTTATGAAACAATGTCTGAATGGAAAAACGAATTCCTGTTATTTTCCCTT 1017  
Qy 321 ArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSer 340  
Db CGTTATCGTTCAAAACCATTTGGTACAGATTCAAGACCCAGAGAACCAAGTCCACACCG 1077

```
QY 341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360
Db 1078 ACTCCAGAACCTAGTCCAACT-----CCGCAACCAAGCTCCAGCAAT 1119
QY 361 ProfileAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380
Db 1120 CCAATTTGATGGGAATTTGGTCAAGAGAGCTGTCGAAAAAGTAGGCGATGGTTATGTCTTT 1179
QY 381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400
Db 1180 GAGGAGATGGAGTTCTCTGTTATATCCAGCCAGAGATCTTTCAGCAGAACAGCAGCA 1239
QY 401 GlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLys 420
Db 1240 GGCATTGATAGCAAACTGCGCAAGCAGGAGAAAGTTTATCTCATAGCTAGGAACTAAGAAA 1299
QY 421 ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440
Db 1300 ACTGACCTCCCATCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1359
QY 441 IleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460
Db 1360 ATTACCAAGATTTACTTGATATTAAGTTCAGCAAGTTGATTTTGGGCTTTGGATAAC 1419
QY 461 LeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeu 480
Db 1420 CTGTTTGAACCACTCAAGGATGCTCAAGTGATATAAGTCAAGTCTAGTGAAGATATTCIT 1479
QY 481 AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr 500
Db 1480 GCCTTTCTTAGCTCCGATTCGTATCAGAACGTTTGTAGGAAAAACCAATGCGCAAAATACC 1539
QY 501 TyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGluAspGly 520
Db 1540 TACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599
QY 521 TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis 540
Db 1600 TATATCTTTGATCTCTGATATACCACTGATGATGATGATGATGATGATGATGATGATGATG 1659
QY 541 MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAla 560
Db 1660 ATGACCCATAGCACTGGATTAAAAAGATAGTTGTCTGCAAGCTGAGAGCGCGAGCC 1719
QY 561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580
Db 1720 CAGGCTTATGCTTAAAGAGAAAGTTTGACCCCTCTTCACAGACCATCAGGATTCAGGA 1779
QY 581 AsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysVal 600
Db 1780 AATACTGAGGCAAAAGGAGCAGAGCTATCTACACCGGTGAAACGAGCTAAGAGGTG 1839
QY 601 ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu 620
Db 1840 CCACCTTGATCGTATGCTTACAACTTCAATATCTCAATATCTGTAAGAGTCAAAACCGTAGT 1899
QY 621 IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeu 640
Db 1900 ATCATACCTCATATGACCAATACCAATACCAATACCAATACCAATACCAATACCAATACCA 1959
QY 641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrVal 660
Db 1960 TATGAGGACCTTAAAGGATATCTCTTGAGGATCTTTTGGGACCTGCTCAAGTACTATGTC 2019
QY 661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680
Db 2020 GAAACATCAAAACGAAAGCTCGCATTCAGATATGTTTGGTAAACCTAGCGACCATGTT 2079
QY 681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLys 700
Db 2080 CAAGAGAAACAAATTTGGTCAAGCTGATACCAATCAACACGAGAAACCAAGCGAGGAGAA 2139
QY 701 ProGlnThrGluLysProGluGluThrProArgGluGluLysProGlnSerGluLys 720
```

```
Db 2140 CCTCAGACAGAAACCTGAGGAAGAAACCCCTCGAAGAGAAACCCCAAGCGAGAAA 2199
QY 721 ProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro 740
Db 2200 CCAAGGTCTCCAAACCAACACAGAGAACCAAGAGAAATCACCAGAGGAATCAGAGAACT 2259
QY 741 GlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuGlyLys 760
Db 2260 CAGGTGAGAGCTGAAAGAGTTGAAGAAAACTGAGAGAGGCTGAAGATTTACTTCGAAA 2319
QY 761 IleGlnAsp 763
Db 2320 ATCCAGGAT 2328
RESULT 4
US-10-412-850-11
; Sequence 11, Application US/10412850
; Publication No. US20040001836A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; FILE OF INVENTION: Motifs
; FILE REFERENCE: 469201-686
; CURRENT APPLICATION NUMBER: US/10/412,850
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 11
; LENGTH: 2531
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-412-850-11
Alignment Scores:
Pred. No.: 0 Length: 2531
Score: 3919.00 Matches: 745
Percent Similarity: 98.03% Conservative: 3
Best Local Similarity: 97.64% Mismatches: 9
Query Match: 97.51% Indels: 6
DB: 16 Gaps: 1
US-09-765-272A-66 (1-763) x US-10-412-850-11 (1-2531)
QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysGluSerAsnArgVal 20
Db 58 TGTTCCTATAGCTTGGACGCTTACCAAGCTGTCAGGATAGAAAGAGCTTAATCGAGTT 117
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 GCTTATATAGTGGTGTATCAGGCTGGTCAAAAGCAGCAAACTTGACACCAAGTGAATC 177
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGTAAGAGGAGGAGGATCAACGCGCAAAATTTATCAAGATTTACGATCAAGGTTAT 237
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTACCTCTCATGAGACCAATATCATCTATATATGCAAGGTTTCCTTATGATGCCATC 297
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 298 ATCAGTGAAGAGCTCTCATGAAGATCCGAATTTATCAGTTGAAGGATTCAGACATGTC 357
QY 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrValTyrLeuLys 120
Db 358 AATGAAATCAAGGCTGGTTTGTGTCATTAAAGTAAACGGTAAATCTATGTTTACCTTAAG 417
```

121 AspAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140  
418 GATGCGCTCATGCCGATATATTCGACAAAGAAGACAGATTAAACGTGAGAGCAGGAA 477  
141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160  
478 CGCAGTCATANTCATAACTCAAGAGCAGATATGCTGTTGCTCAGCAGAGCCCAAGGA 537  
161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180  
538 CGTTATCAACCGATGATGGGTATATCTTCAATGCAATGCTATCATATGAGGACACGGGT 597  
181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200  
598 GATGCTTATATGCTTCTCAGCGGACCATTCATCATATCAATCTCTAAGATGAGTTATCA 657  
201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSer 220  
658 GCTAGCGAGTTAGCTGCTGCAGAGCCTATTGGAATGGGAAGCAGGATCTCGTCCCTCT 717  
221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240  
718 TCAAGTCTAGTTATATGCAATCCAGCTCAACCAAGATTGTTCAGAGAACCACAATCTG 777  
241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260  
778 ACTGTCACCTCCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCCTTTTACGTGAA 837  
261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280  
838 TTGTATGCTTAACCTTATCAGAACCGCATGTGGAATCTGATGCGCTTATTTTCGCCCA 897  
281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300  
898 GCGCAATCACAAGTCGAACCGCAGAGGTGTAGCTGCTCCCTCATGCTAACCATTACCA 957  
301 PheIleProTyrGlnGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeu 320  
958 TTTATCCCTTATGAACAAATGTCTGAATTCGAAACAGAAATGCTGCTATTTATTTCCCTT 1017  
321 ArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSer 340  
1018 CGTTATCGTTCAACCCATTGGTACCGATTCAGATTCAGACAGAGAACCAAGTCCACACCG 1077  
341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360  
1078 ACTCCAGAACCTAGTCCAACT-----CCGCAACCCAGCTCCCAAGCAAT 1119  
361 ProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380  
1120 CCAATTGATGGGAAATTTGGTCAAGAAGCTGTTCCGAAAGTAGCGATGTTATGTTCTTT 1179  
381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400  
1180 GAGGAGAAATCGAGTTCTCGTTATATCCAGCCAGGATCTTTCCAGCAGAACAGCAGCA 1239  
401 GlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLys 420  
1240 GGCATTGATAGCAAACTGGCCACAGCAGGAAGTTTATCTCATAGCTAGGACTAAGAA 1299  
421 ThrAspLeuProSerSerAspArgPheTyrAsnLysAlaTyrAspLeuAlaArg 440  
1300 ACTGACCTCCATCTAGTATCGAGATTTTCAATTAAGGCTTATGACTTACTAGCAAGA 1359  
441 IleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn 460  
1360 ATTCAACCAAGATTTACTTGATTAATAAAGGTGCACAAAGTTGATTTGAGCGCTTGGATTAAC 1419  
461 LeuLeuGluArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeu 480  
1420 CTGTTGGAACGACTCAAGGATGTCTCAAGTGATAAAGTCAAGTTAGTGAAGATATTTCTT 1479

481 AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr 500  
1480 GCCTTCTTAGCTCCGATTCGTCATCCAGACAGCTTTAGGAAACCAAAATGCGCAATATACC 1539  
501 TyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGly 520  
1540 TACACTGATGATGAGATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACAGCAGAGCGGT 1599  
521 TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis 540  
1600 TATATCTTTGATCTCTGATATATAACAGTATGAGGGGATGCTTATGTAATCCACAT 1659  
541 MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAla 560  
1660 ATGACCCATAGCCACTGGATTAAAAAGATAGTTTGTCTGAAGCTCAGAGAGCGGAGCC 1719  
561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580  
1720 CAGGCTTATGCTTAAGAGAAAGTTTGACCCCTCTCTGCACAGACCATCAGGATTCAGGA 1779  
581 AsnThrGluAlaLysGlyValaGluAlaIleTyrAsnArgValLysAlaAlaLysLysVal 600  
1780 AATCTGAGGCAAAAGGAGCAGAGCATATCTACACCGMTGAACGAGTAAAGAGGTG 1839  
601 ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu 620  
1840 CCACCTGATGCTATGCTTCAATCTCTGATAGAGTCAAAACCGTAGTTTA 1899  
621 IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeu 640  
1900 ATCATACCTCATATGACCATTAACCATTAATCAATCTCTGATAGAGTCAAAACCGTAGTT 1959  
641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrVal 660  
1960 TATGAGGCACCTTAAGGGGTATATCTCTGAGGATCTTTTGGCGACTGTCAAGTACTATGTC 2019  
661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680  
2020 GAACATCCAAACGACGTCGCGATTCAGATATGTTTGGTAACTAGCTAGCAGCATGTT 2079  
681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluLys 700  
2080 CAAAGAAACAAATGGTCAAGCTGATCAATCAAAACGAAACCAAGCAGGAGGAAA 2139  
701 ProGlnThrGluLysProGluGluThrProArgGluGluLysProGlnSerGluLys 720  
2140 CCTCAGACAGAAAAACCTGAGAGAGAAACCCCTCGAGAGAGAGAAACCGCAAGCGAGAAA 2199  
721 ProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro 740  
2200 CCAGACTCTCCAAACCAACAGAGGACCAAGAGATCACCAGAGGATCAGAGAACCT 2259  
741 GlnValGluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuGlyLys 760  
2260 CAGGTCGAGACTGAAAGGTTCAAGAAAAACTCAGAGAGGCTCAAGATTTACTTGGAAAA 2319  
761 IleGlnAsp 763  
2320 ATCCAGGAT 2328  
RESULT 5  
US-10-387-783-11  
; Sequence 11, Application US/10387783  
; Publication NO. US2004000531A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; TITLE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-683  
; CURRENT APPLICATION NUMBER: US/10/387,783  
; CURRENT FILING DATE: 2003-03-13



; PRIOR APPLICATION NUMBER: 09/468,656  
 ; PRIOR FILING DATE: 1999-12-21  
 ; PRIOR APPLICATION NUMBER: 60/113,048  
 ; PRIOR FILING DATE: 1998-12-21  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn Ver. 3.0  
 ; SEQ ID NO 11  
 ; LENGTH: 2531  
 ; TYPE: DNA  
 ; ORGANISM: Streptococcus pneumoniae  
 ; US-10-387-783-11

Alignment Scores:	Length:	2531
Pred. No.:	Matches:	745
Score:	Conservative:	3
Percent Similarity:	Mismatches:	9
Best Local Similarity:	Indels:	6
Query Match:	Gaps:	1
DB:		

US-09-765-272A-66 (1-763) x US-10-387-783-11 (1-2531)

QY	1	CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysGlySerAsnArgVal	20
DB	58	TGTTCTATGAGCTTGGAGCTTACCAAGCTGTCAGGATAAGAAAGAGCTTAATCGAGTT	117
QY	21	SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal	40
DB	118	GCTTATATAGATGGTGATCAGGCTGCTCAAAAGGCGAGAAACTTGCACACAGATGAAGTC	177
QY	41	SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr	60
DB	178	AGTAAGAGGCGGGGATCAACCGCAACAAATGTTATCAAGATTACGATCAAGGTTAT	237
QY	61	ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle	80
DB	238	GTCACCTCTATGGAGACCATTAATCAATTAATGCAAGGTTCTTATGATGCCATC	297
QY	81	IleSerGluGluLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal	100
DB	298	ATCAGTGAAGAGCTCTCATGAAGATCCGAATTAATCAAGTTGAAGGATTCAGACATGTC	357
QY	101	AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys	120
DB	358	AATGAATCAAGGCTGGTATGTCATTAAAGGTAACGGTAATACATATGTTACCTTAAG	417
QY	121	AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu	140
DB	418	GATGCGCTCATCGGATATATATTCGACAAAAGAGAGATTAAACGTCAGAACAGGAA	477
QY	141	ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaArgAlaGlnGly	160
DB	478	CGCAGTCATAATCAATACTCAAGAGCAGATAATGCTGTGTCGAGCCAGGCCAAGGA	537
QY	161	ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly	180
DB	538	CGTTATACACCGGATGATGGGTATATCTTCAATGCACTGATATCATTTGAGGACACGGGT	597
QY	181	AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer	200
DB	598	GATGCTTATATCGTTCTTCACGGGACCAATTAATCAATTAATCAATTAATCAATTAATCA	657
QY	201	AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSer	220
DB	658	GCTAGCGAGTTAGTGTGTCAGAGACCTTATGGAATGGAAGGAGGATCTGCTCCTTCT	717
QY	221	SerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu	240
DB	718	TCAAGTTCTAGTTAATGCAATCCAGTCAACCAAGATTGTCAGAGAACCAATCTG	777
QY	241	ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu	260
DB	778	ACTGTCCTCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGAA	837

QY	261	LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro	280
DB	838	TTGTATGCTAAACCCCTTATCAGAACCCCATGTGGAATCTGATGGCCTTATTTTCGACCCA	897
QY	281	AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis	300
DB	898	GCGCAATCACAAGTCGAACCCGAGGCTGAGCTGCTCATGTGTAACCATTAACAC	957
QY	301	PheIleProTyrGluGlnMetSerGluLeuGlyArgIleAlaArgIleIleProLeu	320
DB	958	TTTATCCCTTATGACAAATGCTGAAATTCGAAACGAATTCGCTATTTATTTCCCTT	1017
QY	321	ArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSer	340
DB	1018	CGTTATCGTTCAAAACCATTTGGTCAAGAGCTGTTCCAAAAGTAGCGATGGTATGCTTT	1077
QY	341	ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn	360
DB	1078	ACTCCAGAACCTAGTCCAAGT-----CCGCAACCAAGCTCCAAGCAAT	1119
QY	361	ProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe	380
DB	1120	CCATTTGATGGGAAATTTGGTCAAGAGCTGTTCCAAAAGTAGCGATGGTATGCTTT	1179
QY	381	GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla	400
DB	1180	GAGGAGAATGGAGTTTCTCGTTATATCCAGCCCAAGGATCTTTTCCAGCAAGAACAGCAGCA	1239
QY	401	GlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyValAlaLysLys	420
DB	1240	GGCAATGATGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGGAACCTAAGAAA	1299
QY	421	ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg	440
DB	1300	ACTGACTCTCCATCTAGTATCGAGATTTTACATAAAGGCTTATGACTTACTAGCAAGA	1359
QY	441	IleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsn	460
DB	1360	ATTCAACAAGATTACTTGAATAAAGGTCGACAAAGTTGATTTTGGAGGCTTTGGATAAC	1419
QY	461	LeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***AspIleLeu	480
DB	1420	CTGTGGAACCACTCAAGGATGTCCTCAAGTGATTAAGTCAAGTTAGTGAAGATATCTT	1479
QY	481	AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr	500
DB	1480	GCTTCTTAGCTCCGATTCGTCATCCAGAACGTTTAGGAAAAACCAATGCGCAAAATACC	1539
QY	501	TyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGly	520
DB	1540	TACACTGATGATGAGATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACAGCAAGACGGT	1599
QY	521	TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis	540
DB	1600	TATATCTTTGATCCCTCGTGATATACCGATGATGAGGGGATGCTTATGATCTCCACAT	1659
QY	541	MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAla	560
DB	1660	ATGACCCATAGCCATCGGATTAAAGAGATAGTTTCTGTAAGCTGAGAGCGGAGCC	1719
QY	561	GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly	580
DB	1720	CAGGCTTATGCTTAAAGAGAAAGGTTTACCCCTCTTCGACAGACCATCAGGATTACAGA	1779
QY	581	AsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysVal	600
DB	1780	AATACTAGGCAAAAGGAGCAGAGCTATCTCAACCGMGTGAACAGCAGCTAAGAGGTG	1839
QY	601	ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu	620
DB	1840	CCACTTGATCGTATGCTTCAATCTTCAATATACTGTAGAGTCAAAAACGGTAGTTTA	1899

621 IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeu 640  
1900 ATCAATACCTCATTAAGACCATTTACCATTAACATCAAAATTTGAGTGGTTTGACGAGGCTT 1959  
641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTyrVal 660  
1960 TATGAGGACCTTAAGGGGTATCTCTGAGGATCTTTGGCGACTGTCAAGTACTATGTC 2019  
661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680  
2020 GAACATCCAAACGACGTCGCAATTCAGATAATGGTGTGTAACGCTAGCGACCATGTT 2079  
681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLys 700  
2080 CAAAGAAAACAAATGCTGAGCTGATACCAATCAACGAGAAAACCAAGCGAGGAGAAA 2139  
701 ProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLys 720  
2140 CCTCAGACAGAAAACCTGAGGAGAAACCCCTCGAAGAGAAAACCGCAAGCGAGAAA 2199  
721 ProGluSerProLysProThrGluGluProGluGluSerProGluGluSerGluGluPro 740  
2200 CCAGAGTCTCCAAACCAACAGAGGAAACCAAGAGAAATCACAGAGGAATCAGAGAACCT 2259  
741 GlnValGluThrGluLysValGluGluLysArgGluAlaGluAspLeuGluLys 760  
2260 CAGGTGCGAGCTGAAAGGTTGAAGAAAACCTGAGAGGCTGAGATTACTTTGGAAA 2319  
761 IleGlnAsp 763  
2320 ATCCAGGAT 2328

RESULT 6  
US-09-884-465A-5  
; Sequence 5, Application US/09884465A  
; Publication No. US20030077293A1  
; GENERAL INFORMATION:  
; APPLICANT: Shire Biochem, Inc.  
; APPLICANT: Hamel, Josee  
; APPLICANT: Brodeur, Bernard  
; APPLICANT: Martin, Denis  
; APPLICANT: Charland, Nathalie  
; APPLICANT: Ouellet, Catherine  
; TITLE OF INVENTION: Streptococcus Antigens  
; FILE REFERENCE: 055190-0044  
; CURRENT APPLICATION NUMBER: US/09/884,465A  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: 60/212,683  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 384  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 5  
; LENGTH: 2639  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-09-884-465A-5

Alignment Scores:  
Pred. No.: 0 Length: 2639  
Score: 3665.00 Matches: 700  
Percent Similarity: 92.78% Conservative: 20  
3est Local Similarity: 90.21% Mismatches: 38  
Query Match: 91.19% Indels: 18  
DB: 10 Gaps: 3

US-09-765-272A-66 (1-763) x US-09-884-465A-5 (1-2639)

2y 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20  
Db 171 TGTTCCTATGAACCTGGTGGTCCAGGCTGCTGAGTTAAGAAAGAGTCTAATCGAGTT 230  
2y 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40

Db 231 TCTTATATAGATGGTGGATCAGCTGGTCAAAAGGCAGAGAAAATTTGACACCAAGATGAAGTC 250  
Qy 41 SerLysArgGluGlyIleAsnAlaGluGln\*\*\*ValIleLysIleThrAspGlnGlyTyr 60  
Db 291 AGTAAGAGAGAGGGGATCAACGCCGACAAATTTGTTATCAAGATTACGGATCAAGGTTAT 350  
Qy 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80  
Db 351 GTGACCTCTCATGGAGACCAATATCATTTACTATAATGCGACAGGTTCTTATGATGCATC 410  
Qy 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100  
Db 411 ATCAGTGAAGAACTTCTCATGAAGATCCGAATATACGTTTGAAGGANTCAGACATGTC 470  
Qy 101 AsnGluLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120  
Db 471 AATGAATCAAGGGTGGCTATGTGATTAAAGGTAGACGGAATACTATCTTACCTTAA 530  
Qy 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140  
Db 531 GATCGGCCCATCGGACAATATTCGACAAAGAGAGATTAAACGTCAGAAAGCAGGAA 590  
Qy 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160  
Db 591 CACAGTCAATAATCACTCAAGACGAGATAATGCTTGTCTGCACGACGAGCCCAAGGA 650  
Qy 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180  
Db 651 CGTTATACACGATGATGGGTATATCTTCATGCACTCTGATATCATTTGAGGACACGGGT 710  
Qy 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200  
Db 711 GATGCTTATATCGTTCTCAGCGGACCATTAACATTACATTCCTAAGAATGAGTTATCA 770  
Qy 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSer 220  
Db 771 GCTAGCAGATTAGCTGCTGCAGAGCCTATTGGAATGGGAAGCAGGAGTCTCGTCTCTCT 830  
Qy 221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240  
Db 831 TCAAGTTCTAGTTATATGCAATCCAGTTCACCAAGATTGTCAAGAACCAACATCTG 890  
Qy 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260  
Db 891 ACTGTCACCTCCAACTTAATCATCAAAATCAAGGGGAAAACATTTCAAGCCTTTTACGTA 950  
Qy 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspPro 280  
Db 951 TTGTATGCTAAACCTTATCAGAACGCCATGTAGAATCTGATGGCCTTATTTTCGACCCA 1010  
Qy 281 AlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHis 300  
Db 1011 GCGCAATCACAAGTCGAACCGCAGGTGTAGCTGTCCCTCATGGTAACCATACCAC 1070  
Qy 301 PheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeu 320  
Db 1071 TTTATCCCTTATGACAAATGTCTGAAATGGAAAACGAAATTCCTCTCTATTTATCCCTT 1130  
Qy 321 ArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerProGlnSer 340  
Db 1131 CGTTATCGTTCAACCCATTTGGGTACCAAGATTCAAGACCAAGAACCAACCAAGTCCCAATCG 1190  
Qy 341 ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsn 360  
Db 1191 ACTCGGAACCTAGTCCAAAGTCTGCAACCTGCAACCAAAATCTCAACCGCTCCAAAGCAAT 1250  
Qy 361 ProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPhe 380  
Db 1251 CCAATTTGATGAGAAATTTGTCGAAAGCTGTTGAAAAGTAGAGCGATGTTATGCTTT 1310  
Qy 381 GluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAla 400  
Db 1311 GAGGAGATGAGGTTTCTCGTTATATCCAGCCCAAGGATCTTTTTCAGCAGAAAACAGCAGCA 1370

QY 401 GlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyValAlaLysLys 420  
Db 1371 GGCATTGATAGCAAACTGGCCAGCAGCAAGAAATTATCTCATAGCTAGGAGCTAAGAAA 1430  
QY 421 ThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArg 440  
Db 1431 ACTGACCTCCCATCTAGTATCGAGAAATTTTACAAATAAGGCTTATGACTTATCAGCAAGA 1490  
QY 441 IleHisGlnAspLeuLeuAspAsnLysGlyValArgGlnValAspPheGluAlaLeuAspAsn 460  
Db 1491 ATTCAACAGATTACTTGATAATAAAGGTCGCAAGTTGATTTGATTTGAGTTTGGATAAC 1550  
QY 461 LeuLeuGluArgLeuLysAspVal\*\*\*SexAspLysValLysLeuVal\*\*\*AspileLeu 480  
Db 1551 CTGTTGAAACGACTCAAGATGCTCAAGTATTAAGTCAAGTTAGTGATGATATTCTT 1610  
QY 481 AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThr 500  
Db 1611 GCCTTCTTAGCTCCGATTCTGATCAATCCAGAACGTTTAGGAAACCAATGCGCAAAATTACC 1670  
QY 501 TyrThrAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGly 520  
Db 1671 TACACTGATGATGAGATTCAAGTAGCCAGAGTTGGCAGCAAGTACACCAAGAACCGGT 1730  
QY 521 TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHis 540  
Db 1731 TATATCTTGTGATCTGATATTAACAGTATGAGGGGATGCTATGTAACCTCCACAT 1790  
QY 541 MetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAla 560  
Db 1791 ATGACCCATAGCCACTGGATTAAAGATAGTTTGTCTGAAGCTGAGAGCGCGCAGCC 1850  
QY 561 GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGly 580  
Db 1851 CAGGCTTATGCTAAAGAAAGAGTTTGACCCCTCCCTCGACAGACCACCGAGATTACAGA 1910  
QY 581 AsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysVal 600  
Db 1911 AATACTCAGGCAAAAGGAGCAGAGCTATCTACACCCGCTGGAAGCAGCTAAGAGGTG 1970  
QY 601 ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeu 620  
Db 1971 CCATCTGATCGTATGCCTTCAATCTTCAATATCTGTAGAGTCAAAAACGGTAGTTTA 2030  
QY 621 IleIleProHisTyrAspHisTyrHisAsnLysPheGluTyrPheAspGlnGlyLeu 640  
Db 2031 ATCATACCTCATATTATGACCATTAACCATCAAAATTTGATGTTTGACAGAGCCCTT 2090  
QY 641 TyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrVal 660  
Db 2091 TATGAGGCACCTTAAGGGGTATAGTCTTGAGGATCTTTTGGCAGCTGTCAAGTACTATGTC 2150  
QY 661 GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisVal 680  
Db 2151 GAACATCCAAACGACGCTCCGATTCAGATAATGTTTGGTAAACGCTAGTACCATGTT 2210  
QY 681 GlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLys 700  
Db 2211 CGTAAAAATAAG-----GCAGACCAAGATAGTAAACCTGATGAAGATAAG 2255  
QY 701 ProGlnThrGluLysProGluGluThrProArgGluGluLysProGlnSerGluLys 720  
Db 2256 GAACATGATGAAGTAGTGAGCCAACTCACCTGATCTGATGAAAGAGAAATCAGCT 2315  
QY 721 ProGluSerPro-----LysProThrGluGluProGluGluSerPro 734  
Db 2316 GGTTTAAATCCTTCAGCAGATAATCTTTATAAACCAAGCACTGATACGGAAGACAGAG 2375  
QY 735 GluGluSerGluGlu-----ProGlnValGluThrGluLysVal 747  
Db 2376 GAAAGAGCTGAAGATACCAACAGATGAGGCTGAAATTCCTCAAGTAGAAGATTCTGTATT 2435

QY 748 GluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763  
Db 2436 AACGTAAGATAGCAGATCGGAGCGCTTGTAGAAAAAGTAAACAGAT 2483  
RESULT 7  
US-10-412-862-5  
; Sequence 5, Application US/10412862  
; Publication No. US20040052781A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; TITLE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-685  
; CURRENT FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: US/10/412,862  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 09/468,656  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 5  
; LENGTH: 2531  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-10-412-862-5  
Alignment Scores:  
Pred. No.: 0 Length: 2531  
Score: 3644.50 Matches: 698  
Percent Similarity: 92.41% Conservative: 20  
Best Local Similarity: 89.83% Mismatches: 40  
Query Match: 90.68% Indels: 19  
Gaps: 4  
US-09-765-272A-66 (1-763) x US-10-412-862-5 (1-2531)  
QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyValLysLysGluSerAsnArgVal 20  
Db 58 TGTTCCTATGAATCTGCTGCTCACCAGCTGGTCAGGTTAAGAAAGAGTCTAATCGAGTT 117  
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40  
Db 118 TCTTATATAGTGTGATCAGGCTGCTCAAAAGCGCAAAACTTGACACAGATGAAGTC 177  
QY 41 SerLysArgGluGlyIleAsnAlaGluGln\*\*\*ValIleLysIleThrAspGlnGlyTyr 60  
Db 178 AGTAAGAGGGAGGGGATCAACGCCGAACAAATCGTCATCAAGATTACGGATCAAGGTTAT 237  
QY 61 ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle 80  
Db 238 GTGACCTCTCTGAGAGACCAATATCATCTATATAATGCGAAGGTCCTTATGATGCCATC 297  
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100  
Db 298 ATCAGTGAAGAGCTCTCTCATGAAGATCCGAATATTATCAGTTGAAGGATTTCAGACATTGTC 357  
QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120  
Db 358 AATGAAATCAAGGCTGTTATGTTTATCAAGTAGATGGAATAATATCTATGTTTACTTAAAG 417  
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140  
Db 418 GATGCGAGCTCATCGGATAATATTCGGACAAAGAGAGATTAAACGTCAGACAGGAA 477  
QY 141 ArgSerHisAsnHisAsnSerArgAla---AspAsnAlaValAlaAlaAraGln 159  
Db 478 CACAGTCATTAATCACGGGGGTGTTCTTAACGATCAAGCAGTAGTTGCGAGCCAGAGCCAA 537  
QY 160 GlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThr 179

538 GGACGCTATACACGAGTATGATGGTTATATCTTCAATGCAATCTGATATCATATTGAGGACACG 597  
180 GlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeu 199  
598 GGTGAATCTTATATCTGTTCTTACGGCGACCATTAACATTAACATTAACATTAACATTAACATTA 657  
200 SerAlaSerGluLeuAlaAlaGluAlaTyrTyrAsnGlyLysGlnGlySerArgPro 219  
658 TCAGCTAGCGAGTGTAGTCTGTCAGAAAGCCTATTGGAATGGGAAGCAGGAGTCTCGTCT 717  
220 SerSerSerSerSerTyrAsnAlaAanProAlaGlnProArgLeuSerGluAsnHisAsn 239  
718 TCTTCAAGTCTTATATATATATGCAATCCAGCTCAACCAAGATTGTGAGAAACCAACAT 777  
240 LeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArg 259  
778 CTGACTGTCTCCCACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCCTTTACGT 837  
260 GluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAsp 279  
838 GAATGTATGCTAAACCTTATCAGAACGCCATGTGGAATCTGATGGCCCTTTATTCGAC 897  
280 ProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyr 299  
898 CCAGCGCAATCACAAAGTCCAAACCGCCAGAGGTGTAGCTCTCCCTCATGTGTAACCAATTAC 957  
300 HisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIlePro 319  
958 CACTTTATCCCTTATGAAACAAATGTCTGAATTCGAAACGAAATGTCTGATTTATTTCCC 1017  
320 LeuArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGln 339  
1018 CTTGCTATCGTTCAACCAATGGGTACCAAGATTCAGACCAACCAACCAAGTCCACAA 1077  
340 SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer 359  
1078 TCGACTCCGGAACCTAGTCCAGTCCGCAACCTGCACCAATCTCAACAGCTCCAAAGC 1137  
360 AnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrVal 379  
1138 AATCAATGATGAGAAATGGTCAAGAGCTGTTTCAAGAAAGTAGCGGATGTGTATGTC 1197  
380 PheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAla 399  
1198 TTTGAGGAGAAATGGAGTTCTCTGTTATATCCAGCCAGGATCTTTCAGCAGAAACAGCA 1257  
400 AlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys 419  
1258 GCAGGCATTGATAGCAAACTGGCCAAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAG 1317  
420 LysThrAspLeuProSerSerArgArgGluPheTyrAsnLysAlaTyrAspLeuLeuAla 439  
1318 AAAAAGTACCTCCATCTAGTGTATGAGAAATTTTACAATAGGCTTATGACTTACTAGCA 1377  
440 ArgIleHisGlnAspLeuLeuAspAsnLysGlyValArgGlnValAspPheGluAlaLeuAsp 459  
1378 AGAATTCACCAAGATTTACTTGATTAATAAAGGTCGACAGTTGATTTGAGGCTTGGAT 1437  
460 AsnLeuLeuGluArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIle 479  
1438 AACCTGTTGGAACGACTCAAGGATGTCCTCAAGTATGATAAGTCAAGTTAGTGGATGATTT 1497  
480 LeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIle 499  
1498 CTTGCTCTTCTAGCTCGATTCGTATCCAGAGCTTTAGGAAACCAAAATGCGCAAT 1557  
500 ThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAsp 519  
1558 ACCTACACTGATGATGATTTCAAGTATGACCAAGTTCGAGGAGTACCAACCAAGAGAC 1617  
520 GlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrPro 539  
1618 GGTATATCTTTGATCTCTGATATTAACAGATGATGAGGGGATGCTATGTAACCTCCA 1677

Qy 540 HisMetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaIle 559  
Db 1678 CATATGACCCATAGCCACTGATTAAAAAGATAGTTGTTCTGAAGCTGAGAGAGCGCA 1737  
Qy 560 AlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSer 579  
Db 1738 GCCCAGGCTTATGCTAAAGAGAAAGTTTGACCCCTCTCTGACAGACCATCAGGATTC 1797  
Qy 580 GlyAsnThrGluAlaLysGlyValaGluAlaIleTyrAsnArgValLysAlaAlaLysLys 599  
Db 1798 GGAATACTGAGGCAAAAGAGCAGAGCTATCTCAACCGCTGAAGCAGCTAAGAG 1857  
Qy 600 ValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySer 619  
Db 1858 GTGCCACTGTATGCTGCTTCAATCTTCAATATCTGTAGAAAGTCAAAACCGTAGT 1917  
Qy 620 LeuIleIleProHisTyrAspHisTyrHisIleLysPheGluTrpPheAspGluGly 639  
Db 1918 TTAATCATACCTCATATGACCATTTACCATTAACATCAAAATTTGAGTGGTTTACGAGGC 1977  
Qy 640 LeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyr 659  
Db 1978 CTTTATGAGGCACCTTAAGGGGTATCTCTTGAGGATCTTTTGGCGACTGTCAAGTACTAT 2037  
Qy 660 ValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHis 679  
Db 2038 GTCCGACATCCAAACGAAAGTCCGCAATTCAGATAATGTTTGGTAAACCTAGCGACCAT 2097  
Qy 680 ValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGlu 699  
Db 2098 GTTCGTAATAAATAAG-----GTAGACCAAGACAGACAGTAAACCTGATGAAGAT 2142  
Qy 700 LysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGlu 719  
Db 2143 AAGCAACATGATGAAGTAAGTGAGCCAACTCACCTGAATCTGTATGAAAGAAAGATCAC 2202  
Qy 720 LysProGluSerPro-----LysProThrGluGluProGluGluSer 733  
Db 2203 GCTGGTTTAATCTCTCAGCAGATATCTTTATAACCAAGCACTGATACGGAAGAGACA 2262  
Qy 734 ProGluGluSerGluGlu-----ProGlnValGluThrGluLys 746  
Db 2263 GAGGAAGAAAGCTGAAGATACCAACAGATGAGGCTGAAATTCCTCAAGTAGAGAATTCGT 2322  
Qy 747 ValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763  
Db 2323 ATTAACGCTAAGATAGCAGATGCGGAGGCTTGTCTAGAAAAAGTAACAGAT 2373

## RESULT 8

US-10-412-850-5  
; Sequence 5, Application US/10412850  
; Publication No. US20040001836A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; TITLE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-686  
; CURRENT APPLICATION NUMBER: US/10/412,850  
; CURRENT FILING DATE: 2003-04-14  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 5  
; LENGTH: 2531  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-10-412-850-5

Db	958	CAC	TTTATCCCTTATGAA	CAAAATGCTGTGAATTGGAAAAACGAATGCTGCTATATTATTC	101
Qy	320	Leu	ArgTyrArgSerAsnHis	TrpValProAspSerArgProGluGlnProSerProGln	339
Db	1018	CTT	CGTTATCGTTCAA	ACCATTTGGTACCAGATTCAGACCCAGACCAACCAAGTCCACAA	1077
Qy	340	Ser	ThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer	359	
Db	1078	TCG	ACTCCGGAACTAGTCC	CAAGTCGGCAACCTGCACCAATCCTCAACAGCTCCAAGC	1137
Qy	360	Asn	ProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrVal	379	
Db	1138	AAT	CAATTCATGAGAAATTTGGTCAA	GAAGCTGTTCGAAAGTAGCGCATGTATGTC	1197
Qy	380	Phe	GluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAla	399	
Db	1198	TTT	GAGAGAAATGGAGTTTCTCGTTATATCC	AGCCAGGATCTTTCAGCAGAAACAGCA	1257
Qy	400	Ala	GlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys	419	
Db	1258	CGA	GCATTCATAGCAAACTGGCC	ACGACGAGAAAGTTATCTCATAGCTAGGAGCTAAG	1317
Qy	420	Lys	ThrAspLeuProSerSerAspArgGluPheTyrAsnLysIaLysAspLeuAla	439	
Db	1318	AAA	ACTGACCTCCCATCTAGTGTATCG	AGAAATTTTACAAATGAAGCTTATGATCTTACTAGCA	1377
Qy	440	Arg	IleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAsp	459	
Db	1378	AGA	ATTCCACNAGATTTACTTGATATAA	AGGTTCGACAGTTGATTTTGAGGCTTTGGAT	1437
Qy	460	Asn	LeuLeuGluArgLeuLysAspVal**SerAspLysValLysLeuVal***AspIle	479	
Db	1438	AAC	CTGTGTGAACGACTCAAG	AGTGTCCCAAGTGATAAAGTCAAGTTAGTGGATGATATT	1497
Qy	480	Leu	AlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIle	499	
Db	1498	CTT	CCCTCTCTTAGCTCCGATTCGTCAT	CCAGACGTTTAGSAAACCAAAATGCGCAAAAT	1557
Qy	500	Thr	TyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAsp	519	
Db	1558	ACC	TACACTGATGATGAGATTCAA	GTAGCCUAGTTTGCAGCGAAGTACACACAGAGAC	1617
Qy	520	Gly	TyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrPro	539	
Db	1618	GGT	TATATCTTTGATCCTCGT	GTATATTAACAGTGATCGAGGGATGCTATGTAATCCA	1677
Qy	540	His	MetThrHisSerHisTrpIleLysGlyAspSerLeuSerGluAlaGluArgAlaAla	559	
Db	1678	CAT	ATGACCCATAGCCATCGATTTAAA	AAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGCA	1737
Qy	560	Ala	GlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSer	579	
Db	1738	GCC	CAGGCTTAAGTAAAGAAAGGTTTG	ACCCCTCCTTCGACAGACCATCAGGATTC	1797
Qy	580	Gly	AsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLys	599	
Db	1798	GGA	ATACTCGAGCCAAAGAGCG	AGAGCTATCTTACAAACCGGTGAAGCAGCTAAGAAG	1857
Qy	600	Val	ProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySer	619	
Db	1858	GTG	CACTTGTATCGTATGCGCTTACA	ATCTTCAATATATCTGTAGAAGTCAAAAAACGGTAGT	1917
Qy	620	Leu	IleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGluGly	639	
Db	1918	TTA	ATCATACCTCTATTATGACCATTA	CCATPAACATCAAAATTTGAGTGGTTTGACAGGCG	1977
Qy	640	Leu	TyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyr	659	
Db	1978	CTT	TATGAGGCACCTAAGGGGTAT	CTCTTGAGGATCTTTTGGCGACTGTCAAGTACTAT	2037
Qy	660	Val	GluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHis	679	
Db	2038	GTG	CAACATCCAAACGAACGTCCG	ATTCAGATTAATCGTTTTCGTACGCTAGCAGCAT	2097







141 ArgSerHisAsnHisAsnSerArgAla---AspAsnAlaValAlaAlaAlaArgAlaGln 159  
6079 CACAGTCATATACAGGGGTGGTCTTAACGATCAAGCGAGTAGTTGCAGCCAGAGCCCAA 6138  
160 GlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThr 179  
6139 GGACGCTATACACGGATGATGGTTATATCTTCAATGTCATCTGATATCATTTGAGGACACG 6198  
180 GlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeu 199  
6199 CGTGATGCTTATATCTTCTCACGGCGACCATACCATTTACATTCCTAAGAATGAGTTA 6258  
200 SerAlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgPro 219  
6259 TCAGCTAGCGAGTTAGCTGCTCGAAGAGCCTATGGATGGGAAGCAGGGATCTCGTCT 6318  
220 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsn 239  
6319 TCTTCAAGTCTAGTTATATGCAATCCAGCTCAACCAAGATGTCAGAGAACCACAT 6378  
240 LeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArg 259  
6379 CTGACTGTCTCACTCAACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGT 6438  
260 GluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAsp 279  
6439 GAATGTGATGCTAAACCTTATCAGACGCGATGTGAATCTGAAGCCTTATTTTTCGAC 6498  
280 ProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyr 299  
6499 CAGCGCAATCAACAGTCAACCGCAGAGGTAGTCTGCTCCCTCATGTTAACCATTAC 6558  
300 HisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIlePro 319  
6559 CACTTATCCCTTATGAACAAATGCTGAATTTGAAAAACGAATTTGCTCGTATTATTC 6618  
320 LeuArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGln 339  
6619 CTTTCGTATTCGTTCAACCAATGGTACCAGATTCAAGACCAGAACCAACCAAGTCCCAA 6678  
340 SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer 359  
6679 TCAGCTCCGGAACCTAGTCCAAAGTCCGACCTGACCAATCTCAACAGACTCCAGC 6738  
360 AsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrVal 379  
6739 AATCCATTCATGAGAAATGGTCAAGAAGCTGTTGAAAAAGTAGCGATGGTTATGTC 6798  
380 PheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAla 399  
6799 TTTGAGGAGAAATGGAGTTTCTCGTTATATCCAGCAAGGATCTTTCAGCAGAACACGA 6858  
400 AlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLys 419  
6859 GCAGGCATTTGATAGCAAACTGGCCACGACGAGAAAGTTTATCTCATAGCTAGGACTAAG 6918  
420 LysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuAla 439  
6919 AAACTGACCTCCCATCTAGTATCGAGAAATTTTACAATAAGGCTTATGACTTACTAGCA 6978  
440 ArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAsp 459  
6979 AGAATTCACCAAGATTTACTTTGATAATAAAGTCGACCAAGTTGATTTTTCAGGCTTTGGAT 7038  
460 AsnLeuLeuGluArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIle 479  
7039 AACCTGTGGAAACGACTCAAGATGTCYCAAGTGATAAAGTCAAGTTAGTTAGTGATAT 7098  
480 LeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIle 499  
7099 CTTGCTTCTTAGCTCCCGATTCGTCATTCAGNACGTTTAGGAAACCAATCGCAAT 7158  
500 ThrTyrThrAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGluAsp 519

7159 ACCTACTACTGATGATGAGATTCAAGTAGCCCAAGTTGGCAGGCAAGTACACAAACAGAGAC 7218  
520 GlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrPro 539  
7219 GGTATATATCTTTGATCTCTCGATATACCAAGTATGAGGGGATGCCATATCTAATCCA 7278  
540 HisMetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAla 559  
7279 CATATGACCCCATAGCCACTGCTGATTAAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGCA 7338  
560 AlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSer 579  
7339 GCCCAGGCTTATGCTATAGAGAAAGGTTTGACCTCTCTTCACAGACCATCAGGATTTCA 7398  
580 GlyAsnThrGluAlaLysGlyValaGluAlaIleTyrAsnArgValLysAlaAlaLysLys 599  
7399 GGAATATCTGAGGCAAAAGGAGCAGAGCTATCTACACCGCTGNAAGCAGCAGTAAGAAG 7458  
600 ValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySer 619  
7459 GTGCCACTTGTATCGTATGCTTACATCTTCAATATCTGTAGAACTCAAAACCGTAGT 7518  
620 LeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGly 639  
7519 TTAATCATACCTCATATGACCATTAACCATCAATCAATTTGAGTGGTGTGACGAAGGC 7578  
640 LeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyr 659  
7579 CTTTATGAGGACCACTTAAGGGGTATCTCTTGAGGATCTTTTGGCGACTGTCAAGTACTAT 7638  
660 ValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHis 679  
7639 GTCGAACATCCAAACGACGCTCCGCAATTCAGATAATGGTTTGGTAACGCTAGCAGCAT 7698  
680 ValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGlu 699  
7699 GTTCTGTAATAATAAG-----GTAGACCAACAGACAGTAAACCTGATGAAGAT 7743  
700 LysProGlnThrGluLysProGluGluGluThrProArgGluGluLysProGlnSerGlu 719  
7744 AAGGAACATGATGAAGTAAAGTAGTGAGCAACTCACCTGAAATCTCATGAAAAAGAGATCAC 7803  
720 LysProGluSerPro-----LysProThrGluGluProGluGluSer 733  
7804 GCTGGTTAAATCTCTCAGCAGATAATCTTTATAACCAAGACTGATACGGAAGACACA 7863  
734 ProGluGluSerGluGlu-----ProGlnValGluThrGluLys 746  
7864 GAGGAAGAAGCTGAAGATACCAAGATGAGGCTGAAATCTCTCAAGTAGAGATTTCTGTT 7923  
747 ValGluGluLysLeuArgGluAlaGluAspLeuGlyLysIleGlnAsp 763  
7924 AATTAAGCTAAGATAGCAGATGCGGAGGCTTCTGTAGAAAAAGTAAACAGAT 7974

RESULT 11  
US-09-884-465A-3  
; Sequence 3, Application US/09884465A  
; Publication No. US20030077293A1  
; GENERAL INFORMATION:  
; APPLICANT: Shire Biochem, Inc.  
; APPLICANT: Hamel, Josee  
; APPLICANT: Brodeur, Bernard  
; APPLICANT: Martin, Denis  
; APPLICANT: Charland, Nathalie  
; APPLICANT: Ouellet, Catherine  
; TITLE OF INVENTION: Streptococcus Antigens  
; FILE REFERENCE: 055190-0044  
; CURRENT APPLICATION NUMBER: US/09/884, 465A  
; CURRENT FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: 60/212,683  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 384





```
1969 CATTATGACCAATTACCAATAACATCAAAATTTCAGTGGTTTTCAGCAGGCGCTTTATGAGGCA 2028
644 ProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrThrValGluHisPro 663
2029 CCTAAGGGGTATACCTCTTGGAGGATCTTTGGCGACTGTCAAGTACTATGTGGAACATCCA 2088
664 AsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsn 683
2089 AACGNACTCCGATTCAGATTAATGGTTTGGTAAGCTTAGCGACCATGTTCCAAAGAAC 2148
684 LysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThr 703
2149 AAAAATGGTCAAGCTGATACCAATCAACCGGAAAAACCAAGCGAGGAGAAACCTCAGACA 2208
704 GluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSer 723
2209 GAAAAACCTCGAGGAGAAACCCCTCGAGAAGAGAAACCAAAAGCGAGAAACAGAGTCT 2268
724 ProLysProThrGluGluPro--GluGluSerProGluGluSerGluGluProGlnVal 742
2269 CAAAACCAACAGAGAACCAAGAGAGAAATCAACAGAGGAATCAGAGAACCTCAGGTC 2328
743 GluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuGlyLysIleGln 762
2329 GAGACTGAAAAGGTTGAAGAAAACTGAGAGAGGCTGAAAGATTTACTTGGAAAAATCCAG 2388
763 Asp 763
2389 GAT 2391
RESULT 12
US-09-884-465A-4
Sequence 4, Application US/09884465A
Publication No. US2003007293A1
GENERAL INFORMATION:
APPLICANT: Shire Biochem, Inc.
APPLICANT: Hamel, Josee
APPLICANT: Brodeur, Bernard
APPLICANT: Martin, Denis
APPLICANT: Charland, Nathalie
APPLICANT: Ouellet, Catherine
TITLE OF INVENTION: Streptococcus Antigens
FILE REFERENCE: 055190-0044
CURRENT APPLICATION NUMBER: US/09/884,465A
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 60/212,683
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 384
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 2647
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
US-09-884-465A-4
Alignment Scores:
Pred. No.: 3,896-315 Length: 2647
Score: 3588.50 Matches: 693
Percent Similarity: 91.68% Conservative: 23
Best Local Similarity: 88.73% Mismatches: 44
Query Match: 89.29% Indels: 21
DB: 10 Gaps: 6
US-09-765-272A-66 (1-763) x US-09-884-465A-4 (1-2647)
2Y 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
DB 102 TGTGCTTATGAACCTAGGTTTTCAGTCAAGCT---CAAACTGTAAAGAAAAATAATCGTGT 158
2Y 21 SerTyrIleAspGlyAspGlnAlaGlyGlnIleValAlaGluAsnLeuThrProAspGluVal 40
DB 159 TCCTATATAGATGGAATAACAGCGCAAAAGAACCGAGAAATTTGACTCTCTGATGAGGTT 218
```

```
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
DB 219 AGCAAGCGTGAAGGAATCAACCGCGAACAAATCGTCATCAAGATTACGATCAAGGTTAT 278
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
DB 279 GTGACCTCTCATGGAGACCATTAATCAATTACTATAATGGCAAGGTCCTTATATGCGCATC 338
QY 81 IleSerGluGluLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
DB 339 ATCAGTGAAGAGCTCCATGAAAGATCCGAATTAATCAGTTGAAGGATTCAGACATTGTC 398
QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLysLys 120
DB 399 AATGAATCAAGCGTGGTTATGTCAATTAAGGTAAACGGTAAATACTATGTATTACCTTAAG 458
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
DB 459 GATCAGCTCATGCGGATAATGTCCGTACAAAGAGAAATCAATCGGCAAAACAAAGAA 518
QY 141 ArgSerHisAsnHis-----AsnSerArgAlaAspAsnAlaValAlaAlaArg 157
DB 519 CATAGTCAGCATCGTGAAGGAGGACTTCAGAAACGATGGTCGCTAGCCTTGCACCT 578
QY 158 AlaGlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGlu 177
DB 579 TCACAGGAGCGCTACACACAGATGATGTTTATATCTTCAATGCATCTGATATCATCGAA 638
QY 178 AspThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsn 197
DB 639 GATACGGCGGATGCTATATCGTTCTCATGAGAGATCATTTACATTACATTCTCTTAAGAA 698
QY 198 GluLeuSerAlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLysGlnGly--- 216
DB 699 GAGTTATCAGCTAGCGAGTGGCTGCTGCAAGGCTTCTCTATCTGTCGGGAAATCTG 758
QY 217 -----SerArgProSerSerSerSerSerSerSerSerSerSerSerSerSerSer 226
DB 759 TCAAAATTTAAGAACTATCGCGCGACAAAATAGCGATACACTCCAAGAACAACTGGGTA 818
QY 227 -----AlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThr 243
DB 819 CTTCTCTGTAAGCAATCCAGGAACCTACAAATACTAAACAGCAACCAACAGCAACACTAAC 878
QY 244 ProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAla 263
DB 879 AGT-----CAAGCAAGTCAAAAGTAATGACATTGATGCTCTTGAACACAGCTCTACAA 932
QY 264 LysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIle 283
DB 933 CTGCTTTGAGTCAAGCCATGTAGATCTGATGGCTTATTTTCGACCCAGCGCAATC 992
QY 284 ThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIlePro 303
DB 993 ACAAGTCGAACCGCAGAGGTGTAGTCTGCTCATGTTAAACATTACCACTTTATCCCT 1052
QY 304 TyrGluGlnMetSerGluLeuGlyArgIleAlaArgIleIleProLeuArgTyrArg 323
DB 1053 TATGAACAAATCTCTGAATTTGAAAAACGAATGCTCGTATTAATTTCCCTTCGTATCGT 1112
QY 324 SerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSerThrProGlu 343
DB 1113 TCAACCAATTTGGTACCAGATTCAAGACCAAGAACCAAGTCCACCAACCGACTCCAGAA 1172
QY 344 ProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAsp 363
DB 1173 CTTAGTCCAGTTCGCAACCTGACCAATCTCAACAGCTCCAAGCAATCCAATTGAT 1232
QY 364 GluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsn 383
DB 1233 GAGAAATTTGGTCAAGAGGCTGTTCAAGAAAGTAGCGATGGTTATGTCTTTGAGGAGAA 1292
```

384 GlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAsp 403  
1293 GGAGTTTCTCGTTATATATCCAGCCAGCAAGAAATCTTTTCAGCAGAAAACAGCAGCGCATTCAT 1352  
404 SerLysLeuAlaLysGlnGluSerLysSerHisLysLysLeuGlyAlaLysLysThrAspLeu 423  
1353 AGCAAACTGGCCCAAGCAGGAAAGTTTATCTCATAGCTAGGAGCTAAGAAAACCTGACCTC 1412  
424 ProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaAlaArgIleHisGln 443  
1413 CCATCTAGTGTATCGAGAAATTTCAATAAGCTTATGATCTAGCAAGAAATTCACCAA 1472  
444 AspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGlu 463  
1473 GAUATCTTGATATAAAGGTCGACAAAGTTGATTTGAGGCTTTGGATAAAGCTGTGAA 1532  
464 ArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeuAlaPheLeu 483  
1533 CGACTCAAGGATGTCTCAAGTGAATAAGTCAAGTTAGTGGATGATATTTCTGCGCTTCTTA 1592  
484 AlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAsp 503  
1593 GCTCCGATTCGTATCCAGAACGTTTAGAAAACCAATGGCAATTTACCTACCTCAT 1652  
504 AspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePhe 523  
1653 GATGAGATTCAAGTAGCCAAAGTTGGCAGGCAAGTACACACAGAGACGGTTATATCTTT 1712  
524 AspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHis 543  
1713 GATCCTCGTATATAACAGTGATGAGGGGATGCCTATGTAACCTCCACATATGACCCAT 1772  
544 SerHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyr 563  
1773 AGCCACTGGATTTAAAGAAAGATGTTTGTCTGAGCTGAGAGAGCGGCGCAGCGCTTAT 1832  
564 AlaLysGluLysGlyLeuThrProSerThrAspHisGlnAspSerGlyAsnThrGlu 583  
1833 GCTAAAGAGAAAGTTTGACCCCTCTCGACAGACCATCAGGATTCAGGAATATCTAG 1892  
584 AlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAsp 603  
1893 GCAAAAGCAGCAGAGCTATCTCAACCGCGTGAAGCAGCTAAGAGTGCGCACTTCAT 1952  
604 ArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleLeuPro 623  
1953 CGTATGCTTACAAATCTCAATATCTAGAAAGTCAAAACGGTAGTTTAACTATACCT 2012  
624 HisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeuTyrGluAla 643  
2013 CATTTAGCCATTACCATTAACATCAATTTGAGTGGTTTGAGAGGCTTTATGAGGCA 2072  
644 ProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValGluHisPro 663  
2073 CCTAAGGGGTATATCTTTGAGATCTTTTGGCGACTGTCAAGTACTATGTGCAACATCCA 2132  
664 AsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsn 683  
2133 AACGAACGTCGCGATTCAGATAATGGTTTGGTAAACGCTAGCGACCATGTTCCAAAGAAC 2192  
684 LysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThr 703  
2193 AAAAATGTCAGCTGATACCATCAACCGGAAAAACCCAGCAGGAGGAAACCTCAGACA 2252  
704 GluLysProGluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSer 723  
2253 GAAAAACCTGAGGAGAAACCCCTCGAAGAGAGAAACCAACAAAGCAGAAACAGAGTCT 2312  
724 ProLysProThrGluGluPro---GluGluSerProGluGluSerGluGluProGlnVal 742  
2313 CCAAAACCAACAGAGAGAACCCAGAGAGAAATCACCAGAGGAATCAGAGAACCTCAGGTC 2372  
743 GluThrGluLysValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGln 762

2373 GAGACTGAAAAGGTTGAAGAAAAAAGCTGAGAGGCTGAAGATTACTTGGAAAAATCCAG 2432  
763 Asp 763  
2433 GAT 2435  
RESULT 13  
US-10-412-862-9  
; Sequence 9, Application US/10412862  
; Publication No. US20040052781A1  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Leslie S.  
; APPLICANT: Adamou, John E.  
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
; TITLE OF INVENTION: Motifs  
; FILE REFERENCE: 469201-685  
; CURRENT APPLICATION NUMBER: US/10/412,862  
; CURRENT FILING DATE: 2003-04-14  
; PRIOR APPLICATION NUMBER: 09/468,656  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: 60/113,048  
; PRIOR FILING DATE: 1998-12-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 9  
; LENGTH: 2451  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(2451)  
; OTHER INFORMATION: n = a, c, t or g  
US-10-412-862-9  
Alignment Scores:  
Pred. No.: 6.83e-231 Length: 2451  
Score: 2658.50 Matches: 520  
Percent Similarity: 76.74% Conservative: 77  
Best Local Similarity: 66.84% Mismatches: 132  
Query Match: 66.15% Indels: 49  
DB: 13 Gaps: 8  
US-09-765-272A-66 (1-763) x US-10-412-862-9 (1-2451)  
QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20  
DB 58 TGTTCCTACGAGTTGGGACTGTATCAAGCTAGAACGGTT---AAGGAAAAATAATCGTGT 114  
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40  
DB 115 TCTATATAGATGAGAAAACAAGCGACGCAAAAACGAGAAATTTGACTCTTGATGAGGTT 174  
QY 41 SerLysArgGluGlyIleAsnAlaGluGln\*\*ValIleLysIleThrAspGlnGlyTyr 60  
DB 175 AGCAACGCTGAGGAATCAATCTGAGCAATCTGATCAAGATAACAGACCAAGGCTAT 234  
QY 61 ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle 80  
DB 235 GTCACCTTCACATGGCCGCCACTATCATATTATTACATGTAAGGTTCTTATGACGCTATC 294  
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100  
DB 295 ATCAGTGAAGATTTACTCATGAAGATCCAAACTATAAGCTAAAGATGAGGATTTGTT 354  
QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120  
DB 355 ATGAGGTCAGGGTGGATATGTTATCAAGTAGTGAATAATACATATGTTTACCTTAAG 414  
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluIleLysArgGlnLysGlnGlu 140  
DB 415 GATGCTGCCCGCGGATTAACGTCGCTACAAAAGAGGAAATCAATCGACAAAAACAAGAG 474



; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn Ver. 3.0  
 ; SEQ ID NO 9  
 ; LENGTH: 2451  
 ; TYPE: DNA  
 ; ORGANISM: Streptococcus pneumoniae  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)-(2451)  
 ; OTHER INFORMATION: n = a, c, t or g  
 US-10-412-850-9

## Alignment Scores:

Pred. No.: 6.83e-231 Length: 2451  
 Score: 2658.50 Matches: 520  
 Percent Similarity: 76.74% Conservative: 77  
 Best Local Similarity: 66.84% Mismatches: 132  
 Query Match: 66.15% Indels: 49  
 DB: 16 Gaps: 8

US-09-765-272A-66 (1-763) x US-10-412-850-9 (1-2451)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20  
 Db 58 TGTCTTTACGAGTTGGGACTGTATCAAGCTAGAACGGTT---AAGGAAATAATCGTGT 114  
 QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40  
 Db 115 TCTATATAGTGAACAAAGCAGCAGCAAAACGGAATTTGACTCTCTGATGAGTT 174  
 QY 41 SerLysArgGluGlyIleAsnAlaGluGln\*\*\*ValIleLysIleThrAspGlnGlyTyr 60  
 Db 175 AGCAAGCGTGAAGCAATCAATGCTGAGCAATCGTCATCAAGATAACAGACCAAGCTAT 234  
 QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80  
 Db 235 GTCACTTTCATGCGCAGCACCATATCATTTATACAAATGATGTAAGTTCCCTATGACGCTATC 294  
 QY 81 IleSerGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100  
 Db 295 ATCAGTGAAGAAATTACTCATGAAGATCCAAACTATAAGCTTAAAGATGAGGATATTGTT 354  
 QY 101 AsnGluIleLysGlyLysValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120  
 Db 355 AATGAGTCAAGGTTGATATGTTATCAAGTAGATGGAATAACTATGTTTACCTTAAG 414  
 QY 121 AspAlaIleHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140  
 Db 415 GATGCTGCCACGCGGATAACGTCGTAACAAAGAGGAAATCAATCGACAAACAAAGAG 474  
 QY 141 ArgSerHisAsnHisAsnSer-----ArgAlaAspAsnAlaValAlaAlaAlaArg 157  
 Db 475 CATAGTCAACATCGTGAAGTGGAACTCCAGAAACGATGGTGTGCTTGGCTTGGCAGCT 534  
 QY 158 AlaGlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGlu 177  
 Db 535 TCGCAGACGCTATPACTACAGATGATGTTATCTTTAATGCTTCTGATATCATAGAG 594  
 QY 178 AspThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsn 197  
 Db 595 GATACTGGTATGCTTATATCGTCTCATGAGATCATTAACCATATCATCTTCTTAAGAT 654  
 QY 198 GluLeuSerAlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGly----- 213  
 Db 655 GAGTTATCAGCTAGCGAGTGGCTGCTGAGAAAGCTTCTCTATCTGCTGAGAAATCTG 714  
 QY 214 -----LysGlnGlySerArgProSerSerSerSerSerSerTyrAsn 226  
 Db 715 TCAAAATTCAGAACCTATCGCCGCAAAATAGACATCAACATTCAGAACCAACCTG 774  
 QY 227 AlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThrTyr 246  
 Db 775 CCTTCTGTAGCAATCCAGGAACCTACAAATACTAACAACAGCAACACAGCAACACTAAC 834

QY 247 HisGln---AsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysPro 265  
 Db 835 AGTCAAGCAAGTCAAAAGTAATGACATTGATAGTCTCTTTGAAACAGCTCTACAAAGCTGCT 894  
 QY 266 LeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSer 285  
 Db 895 TTGAGTCAACGACATGTAGAATCTGATGCGCTTGTCTTTGATCCACAAATACAAAGT 954  
 QY 286 ArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGlu 305  
 Db 955 CGAACGCTAGAGGTGTGAGTGCACACGAGATCATTAACCATTCATCCCTTACTCT 1014  
 QY 306 GlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrArgSerAsn 325  
 Db 1015 CAATGTCTGAATTCGAAGAACGATCGCTCGTATTAATTCCTTCCTTCTGTTATCGTTCAAAC 1074  
 QY 326 HisTrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSer 345  
 Db 1075 CATTTGGGTACCAAGATTCAGGCGCAGAACCAACCAAGTCCCAACCGACTCCGGAACCTAGT 1134  
 QY 346 ProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLys 365  
 Db 1135 CCAGCCCGCAGACCTGCACCAATCTTAAA---ATAGACTCAAT-----TCTTCT 1182  
 QY 366 LeuValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyVal 385  
 Db 1193 TTGGTTAGTACGCTGTCGAGAAAGTTGGGGAAGGATATGTATTCGAGAAAGAGGCATC 1242  
 QY 386 SerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys 405  
 Db 1243 TCTCGTTATGCTTTGCGAAAGATTACCATCTGAAACTGTTAAAAATCTTGAAGCAAG 1302  
 QY 406 LeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSer 425  
 Db 1303 TTATCAAAACAAGAGAGTGTTCACACACTTTAACTGCTAAAAAGAAATGTGTCTCT 1362  
 QY 426 SerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeu 445  
 Db 1363 CGTGACCAAGAAATTTATGATAAGCATATATCTGTTAACTGAGGCTCATAAAGCCTTG 1422  
 QY 446 LeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuArgLeu 465  
 Db 1423 TTTGNAATAAGGTCGTAAATTCGATTTCCAGCCTTAGACAAATTTATTAGAAGCCTTG 1482  
 QY 466 LysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeuAlaPheLeuAlaPro 485  
 Db 1483 AATGATGAATCGACTAATAAAGAAATTTGGTAGATGATTTATGGCAATTCCTAGCACCA 1542  
 QY 486 IleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGlu 505  
 Db 1543 ATTACCATCCAGAGCACTTGGCAACCAATCTCAATTCAGTATATCTAGAGCAAG 1602  
 QY 506 IleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAspPro 525  
 Db 1603 GTTTCGTATTGCTCAATTAGCTGATAAGTATACACGCTCAGATGCTTACATTTTGTATGAA 1662  
 QY 526 ArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHis 545  
 Db 1663 CATGATATATCATGATGAAGGAGATGCATATGTAAACGCTCATATGGGCCCATAGTCAC 1722  
 QY 546 TrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAlaLys 565  
 Db 1723 TGGATTGGAAAGATAGCCTTTCTGATAAGGAAAGATTGCAGCTCAAGCTATACATAA 1782  
 QY 566 GluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLys 585  
 Db 1783 GAAAAGGTATCTTACTCTCATCTCCAGACGAGATGTTAAAGCAATCCAACTGGAGAT 1842  
 QY 586 GlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArgMet 605  
 Db 1843 AGTGACGACGACTATTTTACAATCGTGTGAAGGGGAAACGAATTCACACTCGTTCGACTT 1902



366 LeuValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyVal 385  
1183 TTGGTTAGTCAGCTGGTACGAAAGTTGGGAGAGATGATGTTCCGAAGAAAGGCGATC 1242  
386 SerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLys 405  
1243 TCTCGTTATGCTTTCGCAAGATTTACCATCTGAAACTGTTAAATAATCTTGAAGCAG 1302  
406 LeuAlaLysGlnGluSerLysSerHisLysLeuGlyAlaLysLeuThrAspLeuProSer 425  
1303 TTATCAAAACAGAGAGTGTTCACACACTTTAACTGTCTAAAGAAAGAAATGTTCTCT 1362  
426 SerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeu 445  
1363 CGTGACCAAGATTTATGATATAGCATATATCTGTTAACTGAGGCTCATTAAGCCITG 1422  
446 LeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuGluArgLeu 465  
1423 TTTGNAATAAAGGCTGCTAATCTGATTTCCAGGCTTAGACAAATTTATAGAGCGCTG 1482  
466 LysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeuAlaPheLeuAlaPro 485  
1483 AATGATGAATCCACTAATAAAGAAATTTGGTAGATGATTTATTTGGCATTTCTAGACCA 1542  
486 IleArgHisProGluAlaGluLysProAsnAlaGlnIleThrTyrThrAspAspGlu 505  
1543 ATTACCATCCAGAGGCTTGGCAACAATTTCTCAATGAGTATAGTACTGAAGCGNA 1602  
506 IleGlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspPro 525  
1603 GTTCGTATGCTCAATAGCTAGTAAGTATACACGTCAGATGTTTACATTTTGTATGAA 1662  
526 ArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHis 545  
1663 CATGATATAATCAGTGATGAAGAGATGCATATGTAACGCTCATATGGCCATAGTCTAC 1722  
546 TrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGluAlaTyrAlaLys 565  
1723 TGGATTGGAAGAAGTAGCTTTCTGATGAAGAAAGTTGCGAGCTCAAGCCCTATACTAAA 1782  
566 GluLysGlyLeuThrProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLys 585  
1783 GAAAAAGGTATCCTACCTCCATCCAGAGCAGATGTTAAAGCAATCCAACTGGAGAT 1842  
586 GlyAlaGluAlaIleTyrAsnArgValLysAlaAlaLysLysValProLeuAspArgMet 605  
1843 AGTCAGCAGCATTTTACAATCGTGTGAAGGGGAAACGAATTCCTCGACTTTCGACTT 1902  
606 ProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleIleProHisTyr 625  
1903 CCATATATGTTGAGCATACAGTTGAGTTAAACCGTAATTAATGATTAATTTCTCTAAG 1962  
626 AspHisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeuTyrGluAlaProLys 645  
1963 GATCATTAACCAATATTAATTAATTTGTTGTTGATGATCACACATACAAAGCTCCAAAT 2022  
646 GlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisProAsnGlu 665  
2023 GGCTATACCTTGGAGATTTGTTTTCGACGATTAAGTACTAGTGAAACACCTCGACGAA 2082  
666 ArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsn 685  
2083 CGTCCACATTTCAATGATGATGGGCAATGCCAGTGAGCATGTGTAGGCAAGAAAGAC 2142  
686 GlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLys 705  
2143 CACAGTGAAGATCCAAATGAAGACTTCAAGCGGATGAAGAG----- 2184  
706 ProGluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSerProLys 725  
2185 CCAGTAGAGGAACA----- 2199

726 ProThrGluGluProGluGluSerProGluGluSerGluGluProGlnValGluThrGlu 745  
2200 -----CCTGCTGAGCCAGAGTCCCTCAAGTAGAGACTGAA 2235  
746 LysValGluGluLysLeuArgGluAlaGluAspLeuGlyLysIleGlnAsp 763  
2236 AAAGTAGAAGCCCACTCAAGAAAGCAGAGTTTCTTCGAAAGTAGCGAT 2289  
RESULT 16  
US-09-765-272-55  
; Sequence 55 Application US/09765272  
; Patent No. US20020061545A1  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/765,272  
; FILING DATE: 22-Jan-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 55:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2389 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 55:  
US-09-765-272-55

Alignment Scores:  
Pred. No.: 4,31e-230 Length: 2389  
Score: 2649.50 Matches: 519  
Percent Similarity: 76.71% Conservative: 77  
Best Local Similarity: 66.80% Mismatches: 132  
Query Match: 65.92% Indels: 49  
Gaps: 8  
US-09-765-272A-66 (1-763) x US-09-765-272-55 (1-2389)  
QY 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21  
Db 2 TCTTAGAGTTGGAGCTGTATCAAGCTAGACGGTT---AAGGAAATATATCGTGTTC 58  
QY 22 TyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluValSer 41  
Db 59 TATATAGTGAAGAAACAGCGACGCAAAACCGAGAAATTTGACTCTCTGATGAGTTAGC 118  
QY 42 LysArgGluGlyIleAsnAlaGluGln\*\*\*ValIleLysIleThrAspGlnGlyTyrVal 61  
Db 119 AAGCGTGAAGGAATCAATGCTGAGCAAAATCGTCATCAAGATAACAGACCAAGCGCTATGTC 178



62 ThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIlelle 81  
179 ACTTCACATGCGCACCATCATCATATTATTAATGTTAAAGTTCCTTATGACGCTATCATC 238  
82 SerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleValAsn 101  
239 AGTGAAGAATTAATCATGAAGATCCAAATCAATTAAGCTAAAGATGAGGATATGTGAAT 298  
102 GluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLysAsp 121  
299 GAGGTCAGGCTGGATATGTTATCAAGGTAGATGGAATAATCACTATGTTTACCTTAAGGAT 358  
122 AlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGluArg 141  
359 GCTGCCACGCGGATAACCTCGGTACAAAGAGGAATCAATCGACAAACCAAGAGCAT 418  
142 SerHisAsnHisAsnSer-----ArgAlaAspAsnAlaValAlaAlaAlaArgAla 158  
419 AGTCAACATCGTGAAGGTGGAACCTCCAGAAACGATGGTCTGTTCCTTGGCAGCGTTCG 478  
159 GlnGlyArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAsp 178  
479 CAAGAGCGCTACTACTACAGATGATGTTTATGCTTAACTTCTGATATCATAGAGAT 538  
179 ThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGlu 198  
539 ACTGGTGATGCTTATATGCTCTCATGAGATCAITACCATTACATTCCTTAAGATGAG 598  
199 LeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGly----- 213  
599 TTATCAGCTAGCGAGTGGCTGCTGCAGAGAGCTTCTCTATCTGTCGAGGAAATCTGTCA 658  
214 -----LysGlnGlySerArgProSerSerSerSerSerSerTyrAsnAla 227  
659 AATTCAGAAACCTATGCGCGCAAAATAGCGATACATTTCAAGAACAACTGGGTACCT 718  
228 AsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThrTyrHis 247  
719 TCTGTAAGCAATCCAGGAATCAAAATCACTAAACAGCAACCAACCACTAACAGT 778  
248 Gln---AsnGlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyrAlaLysProLeu 266  
779 CAAGCAAGTCAAAAGTAATGACATTTGATGCTCTTGAAACAGCTCTACAAACTGCCCTTG 838  
267 SerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGlnIleThrSerArg 286  
839 AGTCAACGACATGATAGATCTGATGCGCTTGTCTTTGATCCAGCACAAATCACAGTCGA 898  
287 ThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGluGln 306  
899 ACAGCTAGAGGTGTCAGTGCACACGAGATCAATACCATTCATCCCTTTACTCTCAA 958  
307 MetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHis 326  
959 ATGCTCTGAATGGAAGACGAATCGCTGTTATTTCCCTTCTGTTATGTTTCAAAACCAT 1018  
327 TrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSerPro 346  
1019 TGGGTACCAAGATTCAGGCGCAGAACCAACCAAGTCCCAACCGCATCCGGAACCTAGTCCA 1078  
347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366  
1079 GCGCCGCAACCTGCACCAATCTTAA-----ATGACTCAAAAT-----TCTTCTTTG 1126  
367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386  
1127 GTTAGTCAGCTGTCGAAAAGTGGGAAGATATGATTCGAAGAAAGGCGATCTCT 1186  
387 ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406  
1187 CGTTATGCTTTTGCAGAAAGATTTACCATCTGAAATCTTAAATAATCTTGAAGCAAGTTA 1246  
407 AlaLysGlnGluSerLeuSerHisLysLysLeuGlyAlaLysLysThrAspLeuProSerSer 426

1247 TCAAAACAAGAGAGTGTTCACACATTTAACTGTAATAAAGAAATGTTCTCTCTCGT 1306  
427 AspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeu 446  
1307 GACCAAGATTTTAAGTAAGACATATATCTGTTAACTGAGGCTCATAAAGCCCTTGTTT 1366  
447 AspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLys 466  
1367 GNAATAAGGCTCGTAATCTCTGATTTCCAAGCCTTAGACAAATATTATAGACGCTTGAAT 1426  
467 AspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeuAlaPheLeuAlaProIle 486  
1427 GATGAATCGCATTAATAAAGAAATTTGTAGATGATTTATTGTCATTCCTAGCAACCAAT 1486  
487 ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIle 506  
1487 ACCATCCAGAGCAGCTTGGCAACCAATTTCTCAATTTAGTATGATCTAGTACGACGAAT 1546  
507 GlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspProArg 526  
1547 CGTATTGCTCAATTAGCTGATTAAGTATATACAGCTCAGATGGTTACATTTTGTGTAACAT 1606  
527 AspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTrp 546  
1607 GATATAATCAGTATGAAGGAGATGCATATTAACGCTCATATGGCCCATAGTCACTGG 1666  
547 IleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAlaLysGlu 566  
1667 ATTCGAAAAGATAGCCTTTCTGATAAGGAAAAGTTGCGAGCTCAAGCCTTATCTAAAGAA 1726  
567 LysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGly 586  
1727 AAAGTATCTTACCTTCCATCTCCAGACGAGATGTTAAAGCAATCCAACTGGAGTAGT 1786  
587 AlaGluAlaIleTyrAsnArgValLysAlaAlaLysValProLeuAspArgMetPro 606  
1787 GCACGAGCTATTTACATCGTGTGAAGGGGAAAACGAATTCCTCGTCTCGACTTCCA 1846  
607 TyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIleIleProHisTyrAsp 626  
1847 TATATGTTGAGCATACAGTTTCAAGTTTAAACCGTAAATTTGATTTATTTCTCTCATAGGAT 1906  
627 HisTyrHisAsnIleLysPheGluTyrPheAspGluGlyLeuTyrGluAlaProLysGly 646  
1907 CATTACCATAATATTAATTTGCTTGGTTGATGATCACACATACAAAGCTCCAAATGGC 1966  
647 TyrThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisProAsnGluArg 666  
1967 TATACCTTGGAGATTTGTTTGGACGATTAAGTACTAGTAGAACACCTGACGACGT 2026  
667 ProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGly 686  
2027 CCACATTTCTAATGATGGATGGGCAATCCAGTAGCATGTGTAGGCAAGAAAGAACCCAC 2086  
687 GlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLysPro 706  
2087 AGTGAAGATCCAAATAGAACTTCAAGCGGATGAAGAG-----CCA 2128  
707 GluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSerProLysPro 726  
2129 GTAGAGGAACA----- 2140  
727 ThrGluGluProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLys 746  
2141 -----CCTGCTGAGCCAGAGTCCCTCAAGTAGAGACTGAAAAA 2179  
747 ValGluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763  
2180 GTAGAAGCCCAACTCAAAAGAGCAGAACTTTGCTTGGCAAGTAGTAACGAT 2230  
RESULT 17  
US-10-158-844-243





## RESULT 18

US-10-158-844-355

Sequence 355, Application US/10158844

Publication No. US20040029118A1

## GENERAL INFORMATION:

APPLICANT: Kunsch et al.

TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

NUMBER OF SEQUENCES: 391

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

## COMPUTER READABLE FORM:

MEDIUM TYPE: CD-R

COMPUTER: Dell Latitude Pentium 3

OPERATING SYSTEM: Windows 98

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/158,844

FILING DATE: 03-Jun-2002

CLASSIFICATION: &lt;Unknown&gt;

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/961,527

FILING DATE: 1997-10-30

APPLICATION NUMBER: US 60/029,960

FILING DATE: 1996-10-31

ATTORNEY/AGENT INFORMATION:

NAME: Hyman, Mark J.

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB340PID1

SEQUENCE CHARACTERISTICS:

LENGTH: 973 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 355:

US-10-158-844-355

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

2.73e-115

1377.00

98.17%

97.44%

34.26%

13

Length:

Matches:

Conservative:

Mismatch:

Indels:

Gaps:

1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20

155 TGTTCTATGAGCTTGGAGCTTACCGAGTGTCCAGAGTAAAGAAAGCTTATCGATT 214

21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40

215 GCTTATATAGATGGTGTATCAGGCTGTCTAAAGGCGAGAAACCTTGACACAGATGAATC 274

41 SerLysArgGluGlyIleAsnAlaGluGln\*\*\*ValIleLysIleThrAspGlnGlyTyr 60

275 AGTAAAGAGGAGGGGATCAACCGCGAACAATTTGTTATCAAGATTACGGATCAAGGTTAT 334

61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80

335 GTGACCTCTCATGGAGACCATTAATCAATTAATGGCAAGGTTCTTATGATGCATC 394

81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLysLysAspSerAspIleVal 100

395 ATCAGTGAAGAGCTCTCATGAAGATCCGAATTTATCAGTTGAAGATTACAGACATTGTC 454

101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120

Db 455 AATGAATCAAGGGTGGTTATGTCAATTAAGCTAAACCGTAATACTATGTTTACCTTAAG 514

QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140

Db 515 GATGAGCTCATGCGGATATATTCGACAAAAGAGAGATTAAACGTCAGAGCAGGAA 574

QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAraGalaGlnGly 160

Db 575 CGCAGTCATATCATTAATCAAGAGCAGATAATGCTGTGTCGAGCCAGAGCCCAAGGA 634

QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180

Db 635 CGTTATCAACACCGGATGATGCTATATCTTCAATGCAATCTGATATCATTTGAGGACACGGGT 694

QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200

Db 695 GATGCTTATATCGTTCTCCAGCGACCATTAACCATTAATCAATTAAGAAATGAGTTATCA 754

QY 201 AlaserGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLysGlnGlySerArgProSer 220

Db 755 GCTAGCGAGTTAGCTGCTGCAGAGCCCTATTGGAATGGAGACGAGGATCTCGTCTTCT 814

QY 221 SerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu 240

Db 815 TCAAGTTCTAGTTATATGCAAAATCCAGCTCAACCAAGATTGTCAAGAACCAATCTG 874

QY 241 ThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArgGlu 260

Db 875 ACTGTCACTCCACTTATCATCAAAATCAAGGGGAAACATTTCAAGCCTTTTACGTGAA 934

QY 261 LeuTyrAlaLysProLeuSerGluArgHisValGluSer 273

Db 935 TTGTATGCTAA-CCCTTATCAGAACGCCATGTGGATCT 972

## RESULT 19

US-09-884-465A-1

; Sequence 1, Application US/09884465A

; Publication No. US2003007293A1

; GENERAL INFORMATION:

; APPLICANT: Shire Biochem, Inc.

; APPLICANT: Hamel, Josee

; APPLICANT: Brodeur, Bernard

; APPLICANT: Martin, Denis

; APPLICANT: Charland, Nathalie

; APPLICANT: Ouellet, Catherine

; TITLE OF INVENTION: Streptococcus Antigens

; FILE REFERENCE: 055190-0044

; CURRENT APPLICATION NUMBER: US/09/884,465A

; CURRENT FILING DATE: 2001-06-20

; PRIOR APPLICATION NUMBER: 60/212,683

; PRIOR FILING DATE: 2000-06-20

; NUMBER OF SEQ ID NOS: 384

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 3120

; TYPE: DNA

; ORGANISM: Streptococcus pneumoniae

US-09-884-465A-1

## Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

3.85e-101

1228.50

54.31%

40.49%

30.57%

10

Length:

Matches:

Conservative:

Mismatch:

Indels:

Gaps:

US-09-765-272A-66 (1-763) x US-09-884-465A-1 (1-3120)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20

Db 61 TGTCCTATGACTTAACCATTCGTTGCG---CAGGAAATAGGACAAATATCGTGTTC 117

QY	21	SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal	40
Db	118	TCVTATGTGATGCGACCGAGTCAAGTCAAGAAAGTGAAACTTGACACAGACCGAGT	177
QY	41	SerLysArgGluGlyIleAsnAlaGluGln**VallileIysIleThrAspGlnGlyTyr	60
Db	178	AGCCAGAGAAGAGAAATTCAGCGCTGAGCAAAATGTGAATCAAAATTACAGATCAGGGCTAT	237
QY	61	ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle	80
Db	238	GTAACGTCAACGGTGACCACTATCATTAATATGCGAAAGTTCCTTATGATGCCCTC	297
QY	81	IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal	100
Db	298	TTTAGTGAAGAACTCTTGATGAAGAGATCCAAACTATCAACTTAAGACGCTGATATTGTC	357
QY	101	AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys	120
Db	358	AATGAAGTCAAGGTGGTTATATCATCAAGTTCATGGAATAATATATGCTACCTGAAA	417
QY	121	AspAlaIleHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu	140
Db	418	GATGCAGCTCATGCTGATAATGTCGAACATAAGATGAATCAATCAATCGTCAAAACAGAA	477
QY	141	ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValalalalalaArgAlaGlnGly	160
Db	478	CATGTCAAAGAT--AATGAGAAGGTTAACTCTAAATGTTCTGTAGCAAGGTCCTCAGGA	534
QY	161	ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly	180
Db	535	CGATATACGACAATGATGGTTATGCTTTAATCCAGCTGATATATCGAAGATACGGGT	594
QY	181	AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer	200
Db	595	AATGCTTATATCTCTCATGAGGTCACTATCACTATCCAAAACGGATTTATCT	654
QY	201	AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArg	218
Db	655	GCTAGTGAATTACGACGACGATAAAGCACATCTCGCTGGAAAAAATATGCAACGAGTCA	714
QY	219	ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis	238
Db	715	TTAAGCTATTCTTCAACAGCTAGTGACATATACACGCRA-----TCTGTAGCAAAA	765
QY	239	AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeu	258
Db	766	GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAAAATCTCCAGAGTCTTTTG	816
QY	259	ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePhe	278
Db	817	AAGAACTCTATGATTCACTACGCGCCCAAGTTACAGTGAATCAGATGGCCCTGGTCTTT	876
QY	279	AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis	298
Db	877	GACCTGCTAAGATTATCACTCGTACACCAAAATGGAGTTGCGATTCCGATGGCGACCAT	936
QY	299	TyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIle	318
Db	937	TACCACCTTTATTCCTTACAGACAGCTTCTGCTTTAGAAAAAAGATTGCCAATGGTG	996
QY	319	ProLeuArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerPro	338
Db	997	CCATCAGTGGAACTGGTTCTACAGTTTCTACAAATGCAAAACCTAAT-----1044	
QY	339	GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro	358
Db	1045	-----GAAGTAGTGTCTAGCTAGCGACTCTTCA	1074
QY	359	SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr	378
Db	1075	AGCAATCCCTCTCTTTAAACGACAGTAAGAG---CTCTCTCTCAGCATCTGATGGTTAT	1131

379	ValPheGluGluAsnGlyValSerArgTyrIleProAlaIysAspLeuSerAlaGluThr	398
1132	ATTTTAAAT-----CMAAAGATATCGTTCAAGAAACG	1164
399	AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAla	418
1165	GCTACA-----GCTATATTTGAAGACATGGTGATCATTTCCATTAC	1206
419	LysLysThrAspLeuProSerSerArgGluPheTyrAsnLysAlaTyrAspLeuLeu	438
1207	-----ATTCMAAATCAAT-----	1221
439	AlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu	458
1222	---CAAATTTGGGCAACCGACTCTTCCAAACAATAGTCTAGCAACACCTTCTCCATCTTT	1278
459	AspAsnLeuGluArgLeuLysAspVal**SerAspLysValLysLeuVal**Asp	478
1278	-----	1278
479	IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln	498
1279	-----CCAATC-----AATCCAGCA	1293
499	IleThrTyrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGlu	518
1294	ACTTCACATGAGAAACATGAA-----GAA	1317
519	AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr	538
1318	GATGATACGGATTGATGCTAATCGTATTATCGTGAAGATGAATCAGTTTGTCTATG	1377
539	ProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAla	558
1378	AGTCACGGAGACCACAATCATTTATTTCTTCAAGAAGACTTGACAGAAGACCAATTAAG	1437
559	AlaAlaGlnAlaTyrAlaLysGlu-----LysGlyLeuThrProProSer	573
1438	GCTGGCGCAAAACNTTTAGAGGAAGTTAAACCTAGTCAATATGGATTAGATCTTTGTGCA	1497
574	ThrAspHisGlnAsp--SerGlyAsnThr--GluAlaLysGlyAlaGluAlaIleTyr	591
1498	TCTCATGACACGATTTATCCAGGTATGCCAAGAAATGAAGATTTAGATAAA-----	1551
592	AsnArgValLysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGln---	610
1552	-----AAAATCGAAGAAAAAATGCTGGCATTTATGAAACAATATGGTGTCAACCGT	1602
611	-----TyrThrValGluValLysAsnGlySerLeuIleIleProHisTyrAspHisTyr	628
1603	GAAAGATTGTGCGTAATAAGAAAAAATGCGATTATTATCCGCATGGAGATCAACAT	1662
629	His-----AsnIleLysPhe	633
1663	CATGCAGATCCGATTGATGAACATAAACCGTTTGAATTGGTCAATCTTCACAGTAACTAT	1722
634	GluTrpPhe-----AspGluGlyLeu-----TyrGluAlaProLysGlyTyrThrLeu	649
1723	GAACTGTTTAAACCCGACGAAGAGTGTGTAAAAAAGAGGGAATAAAGTTTATCTGGA	1782
650	GluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGluArgProHisSer	669
1783	GAAAGATTAAACGAATGTTGTTAATTTGTTA-----AAA	1815
670	AspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAsp	689
1816	AATAGTACGTTTAAATAATCAAAACCTTACTCTA-----GCCAATGGTCAAAAACGC	1866
690	ThrAsnGlnThrGluLysProSerGluGluLys	700
1867	GTTTCTTTTAGTTTCCCGCTGAATTTGGAGAA	1899

JS-09-884-465A-2  
; Sequence 2, Application US/09884465A  
; Publication No. US20030077293A1  
; GENERAL INFORMATION:

APPLICANT: Shire Biochem, Inc.

APPLICANT: Hamel, Josee

APPLICANT: Brodeur, Bernard

APPLICANT: Martin, Denis

APPLICANT: Charland, Nathalie

APPLICANT: Ouellet, Catherine

**TITLE OF INVENTION: Streptococcus Antigens**

FILE REFERENCE: 055190-0044

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

; CURRENT FILING DATE: 2001-06-20

;; PRIOR APPLICATION NUMBER: 60/212,683

; PRIOR FILING DATE: 2000-06-22

; NUMBER OF SEQ ID NOS: 384

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2

```

; SEQ ID NO 2  
: LENGTH: 50

```

; LENGTH: 50
; TYPE: DNA

```

TYPE: DNA  
ORGANISM:

IS-09-884-4652-2  
ORGANISM: Streptococcus pneumoniae

JS-09-884-465A-2

Alignment Scores:

Alignment Scores:		
Pred. No.:	7.43e-101	Length: 5048
Score:	1228.50	Matches: 296
Percent Similarity:	54.31%	Conservative: 101
Best Local Similarity:	40.49%	Mismatches: 185
Query Match:	30.57%	Indels: 149
DB:	10	Gaps: 24

JS-09-765-272A-66 (1-763) X US-09-884-465A-2 (1-5048)

1	CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValValLysLysGluSerAsnArgVal	20
1837	TGTGCTATGCCTAAACAGCAGCATCGTTGG--CAGGAAATAAAGCAACAATACTCGTGC	1893
21	SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal	40
1894	TCTTATGTGGATGGCAGCAGCTCAAGTCAGAAAGTGAACACTTGCACACAGACCGAGTT	1953
41	SerLysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyr	60
1954	AGCCAGAAAGAGGAATTCAGGCTGAGCAATGTGTAATCAAAATTAACAGATCAGGCGTAT	2013
61	ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle	80
2014	GTAAGCTCACAGCGTGACCACTATCATTAATAATGGGAAAGTCTCTATGATGCGCCT	2073
81	IleSerGluGluLeuLeuMetLysAspProAsnTyrGluLeuLysAspSerAspIleVal	100
2074	TTTAGTGAAGAACTCTTGATGAAGGATCCAACCTATCAACTTAAAGACGCTGATATTGTC	2133
101	AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys	120
2134	ANTGAAGTCAAGGTGGTTATATCATCAAGTCGATGGAAATAATTATGTCTACTCGAA	2193
121	AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu	140
2194	GATCGAGCTCATGCTGATATGTTCCAACTTAAAGATGAAATCAATCGTCAAAACCAAGAA	2253
141	ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaArgAlaGlnGly	160
2254	CATGTCAAAGAT--AATGAGAAGGTTAACTTAATGTTGCTGTAGCAAGGTCTCAGGGA	2310
161	ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly	180
2311	CGATATACGACCAATGATGGTTATGTCTTAAATCCAGCTGATATTTCGAAGATACGGGT	2370
181	AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer	200
2371	AATGCTTATATGCTTCTCATGAGGTCACTATCACTATATCCCAAAAGCGATTATCT	2430

```
QY 559 AlaAlaGlnAlaTyrAlaLysGlu-----LysGlyLeuThrProSer 573
Db 3214 GCTCGCAAAACATTTTACAGCAAGTTAAACTAGCTATAATGATGATTTGTC 3273
QY 574 ThrAspHisGlnAsp---SerGlyAsnThr---GluAlaLysGlyAlaGluAlaIleTyr 591
Db 3274 TCTCATGACAGGATATCCAGGTATNTCCAAAGAAATGAAGATTAGATAAA----- 3327
QY 592 AsnArgValLysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGln--- 610
Db 3328 -----AAAATCGAAGAAAAAATTCGTGCATTATGCAACAATATGGTGTCAACGT 3378
QY 611 -----TyrThrValGluValLysAsnGlySerLeuIlelleProHisTyrAspHisTyr 628
Db 3379 GAAAGTATTGCTGTAATAAAGAAAAAATGCGATTATTTATCCGATGGAGATCACCAT 3438
QY 629 His-----AsnIleLysPhe 633
Db 3439 CATGCAGATCCGATGATGAACATAAACCGGTGGAATTGGTCATTCTCACAGTAACTAT 3498
QY 634 GluTrpPhe-----AspGluGlyLeu-----TyrGluAlaProLysGlyTyrThrLeu 649
Db 3499 GAACGTGTTTAAACCGAAGAGAGGTGCTAAATAAAGAGGGAATAAAGTTTATCTGGA 3558
QY 650 GluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGluArgProHisSer 669
Db 3559 GAAGAAATTAACGAATGTTGTTAAATTGTTA-----AAA 3591
QY 670 AspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAsp 689
Db 3592 AATAGTACGTTTAAATCAAAACCTTACTCA-----GCCAATGGTCAAAACGC 3642
QY 690 ThrAsnGlnThrGluLysProSerGluGluLys 700
Db 3643 GTTTCCTTTAGTTTTCGCGCTCAATTTGAGAAA 3675

RESULT 21
US-10-158-844-192/c
; Sequence 192, Application US/10158844
; Publication No. US20040029118A1
; GENERAL INFORMATION:
; APPLICANT: Kunsch et al.
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESS/SEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude Pentium 3
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/158,844
; FILING DATE: 03-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/961,527
; FILING DATE: 1997-10-30
; APPLICATION NUMBER: US 60/029,960
; FILING DATE: 1996-10-31
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB340PID1
; INFORMATION FOR SEQ ID NO: 192:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6867 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 192:
US-10-158-844-192

Alignment Scores:
Pred. No.: 2,57e-98 Length: 6867
Score: 1202.50 Matches: 294
Percent Similarity: 54.04% Conservative: 101
Best Local Similarity: 40.22% Mismatches: 187
Query Match: 29.92% Indels: 150
DB: 13 Gaps: 23

US-09-765-272a-66 (1-763) x US-10-158-844-192 (1-6867)

QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 6725 TGTGCTATGCTACTAAACCCAGCATCGTTCG---CAGGAAAAATAGGCAATATCGTGTG 6669
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 6668 TCTTATGTGATGGCAGCCAGTCAAGTCAGAAAAAGTGAAAACTTGACACCCAGACCGGTT 6609
QY 41 SerLysArgGluGlyIleAsnAlaGluGln**ValIleLysIleThrAspGlnGlyTyr 60
Db 6608 AGCCAGAAAGAGGAATTTCAGGCTGAGCAAAATGTATCAAAATTACAGATCAGGGGTAT 6549
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 6548 GTAACGTCAACCGTCAGCATCATCATCTACTATAATGGGAAAGTTCCTTATGATGCCCTC 6489
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 6488 TTTAGTGAAGAACTCTTTGATGAAGGATCCAACTATCACTTAAAGACGCTGATATTGTC 6429
QY 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeuLys 120
Db 6428 AATGAAGTCAAGGGTGGTTATATCATCAGTCCATGGAATAATATATATGTCTACTCTGAAA 6369
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
Db 6368 GATGCAAGTCACTGCTGATATGATGTCGAACCTAAGATGAATCAATCTGTCAAAAACAGAA 6309
QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGly 160
Db 6308 CATGTCAAAGAT---AATGAGAAGTTAACTCTAATGTTCTGTAGCAGGTCTCAGGGA 6252
QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 6251 CGATATACGACAAATGATGGTTATGTTTAAATCCAGCTGATATATCGAAGATACGGGT 6192
QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 6191 AATGCTTATATCGTTCTCTCATGAGGTCACTATCATCTACATTCCTCCAAAGCGGATTTATCT 6132
QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLys-----GlnGlySerArg 218
Db 6131 GCTAGTGAATTAGCAGCAGCTAAAGCACATCTCGCTGGAAAAAATATGCAACCGAGTCAG 6072
QY 219 ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238
Db 6071 TTAAGCTATTCTTCAACAGCTAGTAGCAATAACACGCAA-----TCTGTAGCAAAA 6021
QY 239 AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeu 258
Db 6020 GGATCAACTACCAAGCCAGCA-----AATAAATCTGAAAAATCTCCAGAGTCTTTTG 5970
QY 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePhe 278
Db 5969 AAGGAACCTTATGATTCACCTAGCGCCCAACGTTACAGTGAATCAGATGCGCTGGTCTTT 5910
QY 279 AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis 298
Db 5910 AAGGAACCTTATGATTCACCTAGCGCCCAACGTTACAGTGAATCAGATGCGCTGGTCTTT 5910
```

5909 GACCTGCTAAGATTATCATGCTGATACCAAAATGGAGTTCCGATTCGGATCCGATGGACCAT 5850  
299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGlyLysArgIleAlaArgIle 318  
5849 TACCACCTTTATTCCTTACAGCAAGCTTCTGCTGCTAGAAAGAAAGATTGCCAGAATGGTG 5790  
319 ProLeuArgTyrArgSerAsnHisTyrValProAspSerArgProGluGlnProSerPro 338  
5789 CCTATCAGTGGAACTGGTCTTACAGTTTTCACAAATGCAAAACCTAAT----- 5742  
339 GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro 358  
5741 -----GAGTAGTGTCTAGCTAGGAGCTTTTCA 5712  
359 SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378  
5711 AGCAATCCCTCTCTTTAAGCACAAGTAAGAG--CTCTCTTCAGCATCTGATGGTTAT 5655  
379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398  
5654 ATTTTAAAT----- 5622  
399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAla 418  
5621 GCTACA-----GCTTATATTGAAGCATGGTGATCATTTCCATTAC----- 5580  
419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438  
5579 -----ATTCCAAATCAAT----- 5565  
439 AlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458  
5564 ---CAAATTGGGCAACGACTCTTCCAAACAATAGTCTAGCAACCTCTCCATCTCTT 5508  
459 AspAsnLeuLeuGluArgLeuLysAspVal\*\*SerAspLysValLysLeuVal\*\*Asp 478  
5508 ----- 5508  
479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498  
5507 -----CAAATC----- 5493  
499 IleThrTyrThrAspAspGluLeGlnValAlaLysLeuAlaGlyLysTyrThrGlu 518  
5492 ACTTCACATGAGAAACATGAA-----GAA 5469  
519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538  
5468 GATGATACGGATTTGATGCTAATCGTATATATCGTGAAGATGAATCAGGTTTGTCTATG 5409  
539 ProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAla 558  
5408 AGTCACGGAGACCAATCATTTATTTCTCAAGAGGACTTGACAGAGCAAAATTA-- 5350  
559 AlaAlaGlnAlaTyrAlaLysGlu-----LysGlyLeuThrProProSer 573  
5349 GSTGGCAAAACATTAGAGAAAGTTAAACCTAGTCATAATGATTAGATTCTTTGTCA 5290  
574 ThrAspHisGlnAsp-----SerGlyAsnThrGluAlaLysGlyAlaGluAlaLeTyr 591  
5289 TCTCATGACAGGATTTATCCAAAGTAAATGCAAAAGAAATGAAGATTAGATAAA----- 5236  
592 AsnArgValLysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGln--- 610  
5235 -----AAATCGAAGAAATTTGCTGGCATTTATGAAACAATATGATGTGCAACCGT 5185  
611 -----TyrThrValGluValLysAsnGlySerLeuIleIleProHisTyrAspHisTyr 628  
5184 GAAAGTATTGCTGTAATAAGAAAAAATGCGATTATTTATCCGATGAGATCCACAT 5125  
629 His-----AsnIleLysPhe 633  
5124 CATGCAGATCCGATTGATGAACATAAACCGTTGGAAATGGTCAATTCATCTCACAGTAACAT 5065

Qy 634 GluTrpPhe-----AspGluGlyLeu-----TyrGluAlaProLysGlyTyrThrLeu 649  
Db 5064 GAACCTGTTTAAACCCGAGAGAGGAGTGTCTAAAAAGAGGAATAAAGTTTATCTGGA 5005  
Qy 650 GluAspLeuLeuAlaThrValLysTyrTyrValGluHisProAsnGluArgProHisSer 669  
Db 5004 GAAGAAATTAACGAATGTTTAAATTTGTTA-----AAA 4972  
Qy 670 AspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAsp 689  
Db 4971 AATAGTACGTTTAATATCAAAACCTTACTCTA-----GCCAATGGTCAAAACGC 4921  
Qy 690 ThrAsnGlnThrGluLysProSerGluGluLys 700  
Db 4920 GTTCTCTTATGTTTCCGCTCGAATGGAGAAA 4888

## RESULT 22

US-10-158-844-258  
; Sequence 258, Application US/10158844  
; Publication No. US20040029118A1  
; GENERAL INFORMATION:  
; APPLICANT: Kunsch et al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-R  
; COMPUTER: Dell Latitude Pentium 3  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/158,844  
; FILING DATE: 03-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/961,527  
; FILING DATE: 1997-10-30  
; APPLICATION NUMBER: US 60/029,960  
; FILING DATE: 1996-10-31  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hyman, Mark J.  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB340PLD1  
; INFORMATION FOR SEQ ID NO: 258:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1684 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 258:  
US-10-158-844-258

## Alignment Scores:

Pred. No.:	6,36e-99	Length:	1684
Score:	1200.00	Matches:	228
Percent Similarity:	99.56%	Conservative:	0
Best Local Similarity:	99.56%	Mismatches:	1
Query Match:	29.86%	Indels:	1
DB:	13	Gaps:	0

US-09-765-272A-66 (1-763) x US-10-158-844-258 (1-1684)

Qy 535 AlaTyrValThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGlu 554  
Db 3 GCCTATGTAATCCACATATGACCCATAGCCACTGGATTAAAAAAGATAGTTTGTCTGAA 62

QY	555	AlaGluArgAlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProSerThr	574
DB	63	GCTGAGAGCGCGCACCC-CAGCGTTATGCTAAAGAGAAAGGTTGACCCCTCCTTCGACA	121
QY	575	AspHisGlnAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgVal	594
DB	122	GACCATCAGGATTCAGGAAATACTCAGGCNAAAGAGCGAGAGCTATCTACAAACCGCGTG	181
QY	595	LysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGlu	614
DB	182	AAAGCAGCTAAGAAAGTGCCACTTGATCGTAGCGCTTACAATCTTCAATATACTGTAGAA	241
QY	615	VallLysAsnGlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGlu	634
DB	242	GTCAAAACCGTAGTITTAATCAATCACTCAATGACCATTACCAATCAATCAAAATTTGAG	301
QY	635	TrpPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAla	654
DB	302	TGGTTTGACGAGGCGCTTTATGAGGCACCTAAAGGGGTATACTCTTGAGGATCTTTGGCG	361
QY	655	ThrValLysTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGly	674
DB	362	ACTGTCAAGTACTATGTCGAACATCAACACGAAACGTCGCGCATTCAGATAATGGTTTGGT	421
QY	675	AsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGlu	694
DB	422	AACGCTAGCGACCATGTTCAAGAAACAAAAAATGGTCAAGCTGATACCAATCAACCGGAA	481
QY	695	LysProSerGluGluLysProGlnThrGluLysProGluGluGluThrProArgGluGlu	714
DB	482	AAACCAACGCGAGGAAACCTCAGACAGAAAAACCTGAGGAAGAAACCCCTCGAGAAAGAG	541
QY	715	LysProGlnSerGluLysProGluSerProLysProThrGluGluProGluGluSerPro	734
DB	542	AAACCGCAACGCGAGAAACCGAGGTTCTCCAAACCAACAGAGAGAACCGAGAAATCACC	601
QY	735	GluGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArgGluAla	754
DB	602	GAGGAATCAGAGAAACCTCAGTCGAGACTGAAAGGTTTGAGAAAAAACTGAGAGAGCCT	661
QY	755	GluAspLeuLeuGlyLysIleGlnAsp	763
DB	662	GAAGATTACTTGGAAAAAATCCAGAT	688

RESULT 23

RESULT 23  
US-09-769-787-246  
; Sequence 246, Application US/09769787  
; Publication No. US20030091577A1

```

/ IDENTIFICATION NO.: US20090300915 / 7A1
/
/ GENERAL INFORMATION:
/
/ APPLICANT: Microbial Technics Limited
/ APPLICANT: Gilbert, Christophe FG
/ APPLICANT: Hansbro, Philip M
/ TITLE OF INVENTION: Proteins
/ FILE REFERENCE: PWC/P21129WO
/ CURRENT APPLICATION NUMBER: US/09/769,787
/ CURRENT FILING DATE: 2001-01-26
/ PRIOR APPLICATION NUMBER: GB 9816337.1
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: US 60/125164
/ PRIOR FILING DATE: 1999-03-19
/ NUMBER OF SEQ ID NOS: 388
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 246
/ LENGTH: 1455
/ TYPE: DNA
/ ORGANISM: Streptococcus pneumoniae
/ US-09-769-787-246

```

Alignment Scores:		
Pred. No.:	4.6e-96	Length: 1455
Score:	1167.50	Matches: 254
Percent Similarity:	59.50%	Conservative: 78
Best Local Similarity:	45.52%	Mismatches: 123

	Query Match:	29.05%	Indels:	103
	DB:	10	Gaps:	14
 US-09-765-272A-66 (1-763) x US-09-769-787-246 (1-1455)				
QY	1	CysSerTyrGIuLeuGLyArgHisGlnAlaGLyGlnValLysLyeGlusAsnArgVal	20	
Db	61	TGTGGCTATGCATAACCAAGCATCGTTTCG---CAGGAATAATAGCAAATAATCGTGTC	117	
QY	21	SerTyrlleASpGIvAspGInAlaGLyGInLysAlaGIuAsnLeuThrProAspGIuVal	40	
Db	118	TCTATTGTGGATGGCAGCCACTCAAGTCAGAANAAGTGAAACTTGACACCAGCACGGTT	177	
QY	41	SerIysArpGIuGLyIlleAsnAlsGluGln**VallllellyslleThraspGInGLyTyr	60	
Db	178	AGCCAGAAAGGAAGAAATTCAGCGCTCAGCAAAATTGTAAATCAAAAATTACAGATCAGCGCTAT	237	
QY	61	ValThrSerHISglYAspHIstYrHIsfTYRzhenGLyLySvalProTYrAspAlatle	80	
Db	238	GTAACGTCAACGGTGACCATTCAATTACTATTAATGGGAAGTTCCTTATGATGCCCTTC	297	
QY	81	IleSerGIuGIuLeuLeuMetLysAspProAsnTYrGInLeuLYsAspSerAspileVal	100	
Db	298	TTTTAGTGAAGAACCTTCGTGATGAAGATCCAAACTATCAACTTAAAGACGCTGATATTGTC	357	
QY	101	asnGluillelysGlyTYrtyrValillelsValasnglyllystyrtYrValtyrleuLys	120	
Db	358	AATGAAGSTCAAGSGTGTTATCATCAAGGTCGATGGAAATAATTATGTCTACTCGAAA	417	
QY	121	aspAlaalHisAlaspaSenlleArgThrylsGluGluillellysArgGInLYsGInGlu	140	
Db	418	GATCGACTCATGCTGATTAATGTTTCGAACATAAGATGAATCAATCGTCAAAAAACAAGAA	477	
QY	141	ArgSerHISenHISenSerArgAlaspaSnAlaValalaaIAarqAlagInGLy	160	
Db	478	CATGCTCAAAGAT--AATGAGAAGGTTAACTCTAATGTTCGTGTAGCAAGTCTCAGGGA	534	
QY	161	ArgTYrThrrAspAspGLYTyrIllepHeasnAlaserApelilleGluAspThrgly	180	
Db	535	CGATATACGAACAATGATGTTATGTCTTTAAATCCAGCTGATATATCGAAGATACGGGT	594	
QY	181	ASPAlatrileValProHISglYAspHIstYrHIsfTYrILEProLYsasnGluLeuSer	200	
Db	595	AATGCTTATATCGTTCTCTCATGGAGGTCACTATCACTACATTCCTCCAAGAGCATTTACT	654	
QY	201	AlsaSerGIuleulaalaLaGIualatyTrpAsnglyLys-----GinglySerArg	218	
Db	655	GCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGMAAAATAATGCAACCGAGTCAG	714	
QY	219	ProSerSerSerSerTYrAsnAlaasnProalnProArgLeuSerGluasnHis	238	
Db	715	TTAAGCTTATTCTTCAACAGCTAGTAGAACAAFAACACGCCAA-----TCTGTAGCAAAA	765	
QY	239	asnLeuThrValthrProThrTYrHISgLnAspnGLyGluasnIlleSerSerLeuLeu	258	
Db	766	GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAANAATCTCCAGAGTCTTTTG	816	
QY	259	ArgGIuLeuTYralalySproLeuSerGIUArgHISVALGluSerAspGIyleullePhe	278	
Db	817	AAGGAACCTCATGATTCACCTAGCGCCCAACGTTACAGTGAATCAGATGCCTGGCTTTT	876	
QY	279	AspProAlacInlinethrSerArgThrAlaAroGLyValalavalProHISglYasnHis	298	
Db	877	GACCTTGCTTARGATTATCAGTCGTACCAAAATGGAGTTGCGAATCCGCAATGGCGACCAT	936	
QY	299	TYRHISPhelleProTYrGIUGInMetSerGIuLeuGluLysArgilleAlaarGillelle	318	
Db	937	TACCATTATTCTTACAGCAAGCTTTCTGCTTITAGAGAAAAGATGCCAAGATGGTG	996	
QY	319	ProLeuArgTYrArgSerasnHisTrpValProAspSerArgProGIuGInProSerPro	338	
Db	997	CCCTACAGTGGAACTGGTTCTACAGGTTCTTACAAAATGCAAAACCTAAT-----	1044	

```
2Y 339 GlnSerThrProGluProSerProGlnProAlaProAsnProGlnProAlaPro 358
Db 1045 -----GAAGTAGTGTCTAGCTAGGCGAGCTTTTCA 1074
2Y 359 SerAsnProIleAspGluLeuValValGluValArgLysValGlyAspGlyTyr 378
Db 1075 AGCAATCTTCTTCTTAAACGACAGTAAGAG---CTCTCTTCCAGCATCTGATGGTTAT 1131
2Y 379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398
Db 1132 ATTTTAAAT-----CCAAAGATATCGTTGAAGAACG 1164
2Y 399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLysSerHisLysLeuGlyAla 418
Db 1165 GCTACA-----GCTTATATTGTAAGACATGGTGATCATTTCCATTAC----- 1206
2Y 419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438
Db 1207 -----ATTCCAAATCAAAAT----- 1221
2Y 439 AlaArgIleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458
Db 1222 ---CAATTTGGGCAACCGACTCTTCCAAACAATAGTCTAGCAACACCTTCTCCATCTCTT 1278
2Y 459 AspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***Asp 478
Db 1278 ----- 1278
2Y 479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498
Db 1279 -----CCAAATC-----AATCCAGGA 1293
2Y 499 IleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGlu 518
Db 1294 ACTTCACATGAGAAACATGAA-----GAA 1317
2Y 519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
Db 1318 GAUGGATACGGATTGATGCTAATCGTATTATCGCTGAAGATGAATCAGGTTTGTGTCATG 1377
2Y 539 ProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGlu 556
Db 1378 AGTCACGGAGACCAACATATTATTTCTTCAAGAAGGACTTGACAGAAAGACAA 1431
```

## RESULT 24

```
JS-09-769-744A-23
: Sequence 23, Application US/09769744A
: Publication No. US20030134407A1
: GENERAL INFORMATION:
: APPLICANT: Le Page, Richard WF
: APPLICANT: Wells, Jeremy M
: APPLICANT: Hanniffy, Sean B
: APPLICANT: Hansbro, Philip M
: TITLE OF INVENTION: Proteins
: FILE REFERENCE: PWC/P21122WO
: CURRENT APPLICATION NUMBER: US/09/769, 744A
: PRIOR FILING DATE: 2001-01-26
: PRIOR APPLICATION NUMBER: PCT/GB99/02452
: PRIOR FILING DATE: 1999-07-27
: PRIOR APPLICATION NUMBER: GB 9816336.3
: PRIOR FILING DATE: 1998-07-27
: PRIOR APPLICATION NUMBER: US 60/125329
: PRIOR FILING DATE: 1999-03-19
: NUMBER OF SEQ ID NOS: 196
: SOFTWARE: Patent in Ver. 2.1
: SEQ ID NO 23
: LENGTH: 1455
: TYPE: DNA
: ORGANISM: Streptococcus pneumoniae
```

JS-09-769-744A-23

Alignment Scores:

Pred. No.: 4.6e-96 Length: 1455

```
Score: 1167.50 Matches: 254
Percent Similarity: 59.50% Conservative: 78
Best Local Similarity: 45.53% Mismatches: 123
Query Match: 29.05% Indels: 103
DB: 10 Gaps: 14
```

US-09-765-272A-66 (1-763) x US-09-769-744A-23 (1-1455)

```
QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysGluSerAsnArgVal 20
Db 61 TGTGCTTATGCTAAACCGCATCTGTTCCG---CAGGAAATTAAGACATAATTCGTGTC 117
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 TCTTATGTGGATGGCAGCCAGTCAAGTCAGAAAGTGAANAACCTTGACACCAGACCGTT 177
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGCCAGAAAGAGGAATTCAGGCTGAGCAAAATGTAATCAAAATTACAGATCAGGGCTAT 237
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTACGTCACACGGTGACCACTATCATTAATAATGGAAGTTCCTTATGATGCCCTC 297
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLysAspSerAspIleVal 100
Db 298 TTTAGTGAAGAACTCTTGATGAAGGATCCAACTATCACTTAAGACGCTGATATGTC 357
QY 101 AsnGluIleLysGlyGlyTyrValLysValAsnGlyLysTyrTyrValTyrLeuLys 120
Db 358 AATGAAGTCAAGGGTGGTTATATCATCAAGTCGATGGAATAATATTATGTCTACCTGAAA 417
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
Db 418 GATGAGCTCATGCTGATAATGTTTCAACTAAGATGAATCAATCGTCAAAAACAAGAA 477
QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaArgAlaGlnGly 160
Db 478 CATGTCAAAGAT---AATGAGAGGTTAACTCTAATGTTGCTGTAGCAGGCTCAGGGA 534
QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 535 CGATATACGACAAATGATGTTATGCTTTAATCCAGCTGATATTATTCGAAGATACGGGT 594
QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 595 AATGCTTATATCGTTCTTCATGGAGGTCATCATCACTACATCCCAAAACGATTTATCT 654
QY 201 AlaSerGluLeuAlaAlaGluAlaTyrTyrAsnGlyLys---GlnGlySerArg 218
Db 655 GCTAGTGAATTAAGCAGCAGCTAAACGACATCTGGCTGGAAAAAATAATGCAACCGAGTCAG 714
QY 219 ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238
Db 715 TTAAGCTATTCTTCAACAGCTAGTGACAAATACACGCAA-----TCTGTAGCAAAA 765
QY 239 AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeu 258
Db 766 GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAAAAATCTCCAGAGTCTTTTG 816
QY 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePhe 278
Db 817 AAGAACTCTATGATTCACCTAGGCCCAAGCTTACAGTGAATCAGATGGCTGCTGCTTT 876
QY 279 AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis 298
Db 877 GACCTGCTAAGATTATCAGTCGACCAAAATGAGTTGGATTCCGATCCGACCAT 936
QY 299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIle 318
Db 937 TACCACCTTATCTCTTACAGCAAGCTTTCTGCTTAGAAGAAAGATTGCCAGATGGTG 996
QY 319 ProLeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerPro 338
```



```

Db 997 CCTATCAGTGGAACTGGTTCTACAGTTTCTCAAAATGCAAAACCTTAAT----- 1044
Qy 339 GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro 358
Db 1045 -----GAGTAGTGTCTAGTCTAGCGAGTCTTTCA 1074
Qy 359 SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378
Db 1075 AGCAATCCTTCTTTTAAACGAAAGTAGGAG---CTCTCTTACGACTCTGATGGTTAT 1131
Qy 379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398
Db 1132 AATTTTAAT-----CCAAAAGATATCGTTGAAGAAACG 1164
Qy 399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLysSerHisLysLeuGlyAla 418
Db 1165 GCTACA-----GCITATATTGACACATCGGTGATCATTTCCATTAC----- 1206
Qy 419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438
Db 1207 -----ATTCAAAATCAAAT----- 1221
Qy 439 AlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458
Db 1222 ---CAAATGGCAACCGACTCTTCCAAACAATAGTCTAGCAACACCTTCTCCATCTCTT 1278
Qy 459 AspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***Asp 478
Db 1278 ----- 1278
Qy 479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498
Db 1279 -----CCAATC-----AATCCAGGA 1293
Qy 499 IleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGlu 518
Db 1294 ACTTCACATGAGAAACATGAA-----GAA 1317
Qy 519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
Db 1318 GATGATACGGAATTGATGCTAATCGTATTATTCGCTGAAGATGATCAGGTTTGTTCATG 1377
Qy 539 ProHisMetThrHisSerHisTyrIleLysLysAspSerLysSerLeuSerGluAlaGlu 556
Db 1378 AGTCACGAGAGACCAATCAATATTCTTCAAGAGGACTTGACAGAGAGACAA 1431
RESULT 25
US-10-412-862-7
; Sequence 7, Application US/10412862
; Publication No. US20040052781A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamon, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-685
; CURRENT APPLICATION NUMBER: US/10/412,862
; PRIOR FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-412-862-7
Alignment Scores:

```

```

Pred. No.: 4,6e-96 Length: 1455
Score: 1167.50 Matches: 254
Percent Similarity: 59.50% Conservative: 78
Best Local Similarity: 45.52% Mismatches: 123
Query Match: 29.05% Indels: 103
DB: 13 Gaps: 14

```

US-09-765-272a-66 (1-763) x US-10-412-862-7 (1-1455)

```

Qy 1 CysSerThrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 61 TGTGCTTATCCATTAACACGACATCGTTTCG---CAGGAAATTAAGCAATAATATCGTGTCT 117
Qy 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 TCTTATGTGATGGCAGCCAGTCAGTCAGTCAAGTCAAGAAAGTGAAACCTTGACACCAAGCCAGTT 177
Qy 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGCCAGAAAGAGGAATTCAGGCTGAGCAAAATTGTAATCAAAATTACAGATCAGGGCTAT 237
Qy 61 ValThrSerHisGlyAspHisTyrHisTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTAACGTCAACCGGTGACCACTATCTATCTATTAATGGGAAAGTTCTTATGATGCCCTC 297
Qy 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 298 TTTAGTGAGAACTCTTGTGAGAGGATCCAACTATCACTTAAAGCGGTGATTTGTC 357
Qy 101 AsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLys 120
Db 358 AATGAAGTCAAGGCTGTTATATCATCATCAAGGTGATGAGAAATATTATGTTCTCTCTGAAA 417
Qy 121 AspAlaHisAlaAspAsnIleArgThrLysGluGluLysArgGlnLysGlnGlu 140
Db 418 GATGACGCTCATGCTGATATGTTGCAACTAAGATGAAATCAATCTGCAAAAACAAGAA 477
Qy 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAlaAlaGlnGly 160
Db 478 CATGTCAAAGAT---AATGAGAAGTTTAATCTTAATGTTCTGTAGCAAGCTCTCAGGA 534
Qy 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 535 CGATATACACAAATGATGTTATGCTTTAATCCAGCTCATATTATCGAAGATACGGGT 594
Qy 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 595 AATGCTTATATCGTTCTCTCATGAGGTCACTATCACTACATTCCTCCAAAAGCGATTTATCT 654
Qy 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTyrAsnGlyLys-----GlnGlySerArg 218
Db 655 GCTAGTGATTTAGCAGCAGCTAAAGCACATCTGGTGGAAAAAATAATGCAACCGAGTCAG 714
Qy 219 ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238
Db 715 TTAAGCTATTCTTCAACAGCTAGTGACATAAACAACGCA-----TCTGTAGCAAAA 765
Qy 239 AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeu 258
Db 766 GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAAAATCTCCAGAGTCTTTTG 816
Qy 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePhe 278
Db 817 AAGGAATCTATGATTCACCTAGCGCCCAACGGTTCACAGTGAATCAGATGCCCTGCTTT 876
Qy 279 AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis 298
Db 877 GACCTGCTAAGTATATCATGTCACCAATGAGTTGCGATTCCGATGCGGACCAT 936
Qy 299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGlyLysArgIleAlaArgIle 318
Db 937 TACCACCTTATTCCTTACAGCAAGCTTCTGCTTAGAAGAAAGATTCAGAAATCGTG 996

```

319 ProLeuArgTyrArgSerAsnHisTirValProAspSerArgProGluGlnProSerPro 338  
|||::: :::  
997 CCTATCAGTGGAACTGGTCTTACAGTTCTACAAATGCAAACTAAT----- 1044  
339 GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro 358  
::: :::  
1045 -----GAAGTAGTGCTAGTCTAGGACGCTCTTTCA 1074  
359 SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378  
|||::: |||:::  
1075 AGCAATCCTTCTCTTAAACGACAACTAAGGAG--CTCTCTTCAGCATCTGATGGTAT 1131  
379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398  
||||::: |||:::  
1132 ATTTTAAAT-----CCAAAGATATCGTTGAAGAAACG 1164  
399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnLysSerLeuSerHisLysLeuGlyAla 418  
::: ::: |||:::  
1165 GGTACA-----GCTTATATTGTAAACATGGTGATCATTTCCATTAC----- 1206  
419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438  
::: |||:::  
1207 -----ATTCAAAATCAAT----- 1221  
439 AlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458  
|||::: |||:::  
1222 ---CAAATTTGGGCAACCGACTCTTCCAAACAATAGTCTAGCAACACCTTCTCCATCTCT 1278  
459 AspAsnLeuLeuGluArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*Asp 478  
1278 ----- 1278  
479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498  
|||::: |||:::  
1279 -----CCAATC-----AATCCAGGA 1293  
499 IleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGlu 518  
|||::: |||:::  
1294 ACTTCACATGAGAAACATGAA-----GAA 1317  
519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluLysAspAlaTyrValThr 538  
|||::: |||:::  
1318 GATGATACGGATTGATGCTAATCGTATTATTCGCTGAAGATGAATCAGGTTTGTGTCATG 1377  
539 ProHisMetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGlu 556  
|||::: |||:::  
1378 AGTCACGAGGACCACAATCATTTATTTCTCAAGAGGACTTCACAGAGAGCAA 1431

RESULT 26  
IS-10-412-850-7  
Sequence 7, Application US/10412850  
Publication No. US20040001836A1  
GENERAL INFORMATION:  
APPLICANT: Johnson, Leslie S.  
APPLICANT: Adamou, John E.  
TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus  
TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural  
TITLE OF INVENTION: Motifs  
FILE REFERENCE: 469201-696  
CURRENT APPLICATION NUMBER: US/10/412,850  
CURRENT FILING DATE: 2003-04-14  
PRIOR APPLICATION NUMBER: 09/468,656  
PRIOR FILING DATE: 1999-12-21  
PRIOR APPLICATION NUMBER: 60/113,048  
PRIOR FILING DATE: 1998-12-21  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 3.0  
SEQ ID NO 7  
LENGTH: 1455  
TYPE: DNA  
ORGANISM: Streptococcus pneumoniae  
IS-10-412-850-7

```

QY 319 ProLeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerPro 338
Db 997 CCTATCATGGAACCTGGTTCTACAGCTTTCTACAAATGCAAAACCTAAAT----- 1044
QY 339 GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro 358
Db 1045 -----GAAAGTAGTGCTAGTCTAGGAGAGCTTTCA 1074
QY 359 SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378
Db 1075 AGCAATCTCTCTTTAAACGACAAAGTAGGAG---CTCTCTTCAGCATCTCATGGTTAT 1131
QY 379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398
Db 1132 ATTTTAAAT-----CCAAAAGATATCGTTGAAGAAACG 1164
QY 399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLysSerHisLysLysGlyAla 418
Db 1165 GCTACA-----GCTTATATTCTAAGACATGCTGATCATTCATAC----- 1206
QY 419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438
Db 1207 -----ATTCCAAATCAAT----- 1221
QY 439 AlaArgIleHisGlnAspLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458
Db 1222 ---CAAATTTGGCAACCGACTCTTCCAAACATAGTCTAGCAACACCTTCTCCATCTCTT 1278
QY 459 AspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal***Asp 478
Db 1278 ----- 1278
QY 479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498
Db 1279 -----CCAATC-----AATCCAGGA 1293
QY 499 IleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGlu 518
Db 1294 ACTTCACATGAGAACATGAA-----GAA 1317
QY 519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538
Db 1318 GATGGATACGGATTGTGCTAATCGTATTATCGCTGAAGATGAATCAGGTTTGTCTATG 1377
QY 539 ProHisMetThrHisSerHisTrpIleLysLysAspSerLysSerGluAlaGlu 556
Db 1378 AGTCACGGAGACCAATCATATTCTTCAAGAAGGACTTGACAGAGAGCAA 1431
RESULT 27
US-10-387-783-7
; Sequence 7, Application US/10387783
; Publication No. US2004000531A1
; GENERAL INFORMATION:
; APPLICANT: Johnson, Leslie S.
; APPLICANT: Adamou, John E.
; TITLE OF INVENTION: Vaccine Compositions Comprising Streptococcus
; TITLE OF INVENTION: Pneumoniae Polypeptides Having Selected Structural
; TITLE OF INVENTION: Motifs
; FILE REFERENCE: 469201-683
; CURRENT APPLICATION NUMBER: US/10/387,783
; CURRENT FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: 09/468,656
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/113,048
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 7
; LENGTH: 1455
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-387-783-7

```

```

Alignment Scores:
Pred. No.: 4,6e-36 Length: 1455
Score: 1167.50 Matches: 254
Percent Similarity: 59.50% Conservative: 78
Best Local Similarity: 45.52% Mismatches: 123
Query Match: 29.05% Indels: 103
DB: 16 Gaps: 14

US-09-765-272A-66 (1-763) x US-10-387-783-7 (1-1455)
QY 1 CysSerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgVal 20
Db 61 TGTGCGCTATGCACTAAACCCAGCATCGTTGC---CAGGAAATTAAGCAATAATTCGTGTC 117
QY 21 SerTyrIleAspGlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluVal 40
Db 118 TCTTATGTGGTGGCCGCCAGTCAAGTCAGAAAAGTGAANAACCTGCACACCCAGCAGTT 177
QY 41 SerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGlyTyr 60
Db 178 AGCCAGAAAGAGGAATTCAGGCTGAGCAATTTGTAATCAAAATTACAGATCAGGCTAT 237
QY 61 ValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIle 80
Db 238 GTAACGCTCACCGGTGACCACTATCATTAATAATGGAAAGTTCTCTTATGATGCCCTC 297
QY 81 IleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIleVal 100
Db 298 TTTAGTGAAGAACTCTTGATGAAGGATCCAACTATCAACTTAAAGACGCTGATATTGTC 357
QY 101 AsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrValLys 120
Db 358 AATGAAGTCAAGGTGGTTATATCATCATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 417
QY 121 AspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnGlu 140
Db 418 GATGACGCTCATGCTGATTAATGTTCAACTTAAAGATCAAAATCAATCGTCAAAAAACAAGAA 477
QY 141 ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValAlaAlaAalaArgAlaGlnGly 160
Db 478 CATGTCAAAGAT---AATGAGAAGTTAACTTAATGTTGCTAGCAAGGCTCTCAGGGA 534
QY 161 ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly 180
Db 535 CGATATACGACAAATGATGCTGATGCTTTAATCCAGCTGATATTATTCGAAGATACGGGT 594
QY 181 AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer 200
Db 595 AATGCTTATATCGTCTCTCATGAGGTCACTATCACTTACATTCCTCAAAAGCGATTATCT 654
QY 201 AlaSerGluLeuAlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArg 218
Db 655 GCTAGTGAATTAGCAGCAGCTAAGACACATCTGGTGGAAAAAATATGCAACCCAGGTCAG 714
QY 219 ProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHis 238
Db 715 TTAAGCTATTCTTCAACAGCTAGTAGCAATAACACGCAA-----TCTGTAGCAAAA 765
QY 239 AsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyAsnIleSerSerLeuLeu 258
Db 766 GGATCAACTAGCAAGCCAGCA-----AATAAATCTGAAAATCTCCAGAGTCTTTTG 816
QY 259 ArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePhe 278
Db 817 AAGGAACCTCTATGATTCACTACGCGCCCAACGTTACAGTGAATCAGATGGCCTGCTTT 876
QY 279 AspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHis 298
Db 877 GACCTGCTAAGATTATCATGCTACCAAAATGGAGTTGCGATTCCGATGGCGCAT 936
QY 299 TyrHisPheIleProTyrGluGlnMetSerGluLeuGlyLysArgIleAlaArgIle 318
Db 936 TTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTAT

```

937 TACCACTTTATCTCTTACAGCAAGCTTTCTGCTTAGAAGAAAGATTGCCAATGGTG 996  
319 ProLeuArgTyrArgSerAsnHisTrpValProAspSerArgProGluGlnProSerPro 338  
997 CCTATCAGTGAAGTGGTCTTACAGTCTTACAAATGCAAAACCTAAT----- 1044  
339 GlnSerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaPro 358  
1045 -----GAAGTAGTGTCTAGTCTAGGCACTTTTCA 1074  
359 SerAsnProIleAspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyr 378  
1075 AGCAATCTCTCTTTACAGCAAGTAGAG--CTCTCTCAGCATCTGATGGTAT 1131  
379 ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThr 398  
1132 ATTTTAAAT-----CCAAAGATATCGTTGAAGAACG 1164  
399 AlaAlaGlyIleAspSerLysLeuAlaLysGlnLysSerHisLysLeuGlyAla 418  
1165 GCTACA-----GCTTATATTGTAAGACATGGTGATCATTTCCATTAC----- 1206  
419 LysLysThrAspLeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeu 438  
1207 -----ATTCCAAATCAAAAT----- 1221  
439 AlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeu 458  
1222 ---CAAATGGGCAACCGACTCTTCCAAACATAGTCTAGCACACCTTCTCCATCTCT 1278  
459 AspAsnLeuLeuGluArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*Asp 478  
1278 ----- 1278  
479 IleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGln 498  
1279 -----CCAAATC-----AATCCAGGA 1293  
499 IleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGlu 518  
1294 ACTTCACATGAGAAACATGAA-----GAA 1317  
519 AspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThr 538  
1318 GATGATACGGATTTGATGCTAATCGTATTATTCGCTGAAGATGAATCAGGTTTGTGATG 1377  
539 ProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGlu 556  
1378 AGTCACGGAGACCAATCATTTCTTCAAGAGGACTTGACAGAGAGCAA 1431

## RESULT 28

US-09-765-272-181  
Sequence 181, Application US/09765272  
Patent No. US20020061545A1  
GENERAL INFORMATION:  
APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,272  
FILING DATE: 22-Jan-2001

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/961,083  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8512  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 181:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1342 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 181:  
US-09-765-272-181

Alignment Scores:  
Pred. No.: 9,36e-94 Length: 1342  
Score: 1141.50 Matches: 250  
Percent Similarity: 59.41% Conservative: 75  
Best Local Similarity: 45.70% Mismatches: 119  
Query Match: 28.40% Indels: 103  
DB: 9 Gaps: 14

US-09-765-272a-66 (1-763) x US-09-765-272-181 (1-1342)

QY 5 LeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSerTyrIleAsp 24  
DB 2 CTAACCAAGCATCGTTCG---CAGGAAATTAAGGACAATAATCGTCTCTTATGTGGAT 58  
QY 25 GlyAspGlnAlaGlyGlnLysAlaGluAsnLeuThrProAspGluValSerLysArgGlu 44  
DB 59 GGCAGCCAGTCAAGTCAAGAAAGTGAACCACTTGACACAGACCCAGGTTAGCCAGAAAGAA 118  
QY 45 GlyIleAsnAlaGluGln\*\*\*ValIleLysIleThrAspGlnGlyTyrValThrSerHis 64  
DB 119 GGAATTCAGGCTGAGCAAAATTTGTAATCAAAATTAACAGATCAGGGCTATGTACGTCACAC 178  
QY 65 GlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAlaIleIleSerGluGlu 84  
DB 179 GGTGACCATCATCTACTATAATGGGAAAGTCTCTTATGATGCCCTCTTTAGTGAAGAA 238  
QY 85 LeuLeuMetLysAspProAsnTyrGlnLysLysAspSerAspIleValAsnGluLys 104  
DB 239 CTCTTGATGAAGGATCCAAACTATCACTTAAGACGCTGATATTGTCAATGAAGTCAAG 298  
QY 105 GlyGlyTyrValIleLysValAsnGlyLysTyrValTyrLeuLysAspAlaAlaHis 124  
DB 299 GGTGGTTATATCATCAAGTGTGAGGAAATATTATGTTCTACCTGAAAGATCCAGCTCAT 358  
QY 125 AlaAspAsnIleArgThrLysGluGluIleLysArgGlnLysGlnLysGlnLysArgSerHisAsn 144  
DB 359 GCTCATATGTTCGAATCAAGATGAATCAATCGTCAAAACAGAACATGTCAGAGAT 418  
QY 145 HisAsnSerArgAlaAspAsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrThr 164  
DB 419 ---AATGAGAAGGTTAACTCTAATGTTGCTGTAGCAAGGCTCTCAGGACGATATACGACA 475  
QY 165 AspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIle 184  
DB 476 AATGATGGTTATGTCTTTAATCCAGCTGATATATTATCGAAGATACGGGTAATGCTTATATC 535  
QY 185 ValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeu 204  
DB 536 GTTCTCATGGAGGTCACTACTACATACATTCCTCCAAAGCGATTATCTGCTAGTAGAATTA 595  
QY 205 AlaAlaAlaGluAlaTyrTrpAsnGlyLys-----GlnGlySerArgProSerSer 222  
DB 596 GCAGCAGCTAAAGCACATCTCGCTGGGAAAAAATAATGCAACCGAGTCAGTTAAGCTATTCT 655

QY 223 SerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrVal 242  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 243 ThrProThrTyrHisGlnAsnGlnGlyGluAsnLeuSerSerLeuLeuArgGluLeuTyr 262  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 263 AlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGln 282  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 283 IleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPhe 302  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 303 ProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleLeuProLeuArgTyr 322  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 323 ArgSerAsnHisTyrValProAspSerArgProGluGlnProSerProGlnSerThrPro 342  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 343 GluProSerProSerProGlnProAlaProAsnProGlnProAlaProSerAsnProIle 362  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 363 AspGluLysLeuValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGlu 382  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 383 AsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIle 402  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 403 AspSerLysLeuAlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAsp 422  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 423 LeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHis 442  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 443 GlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeu 462  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 463 GluArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeuAlaPhe 482  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 483 LeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThr 502  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 503 AspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIle 522  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 523 PheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThr 542  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 543 HisSerHisThrPheLysLys 549  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706  
QY 549 HisSerHisThrPheLysLys 549  
Db TCAACAGCTAGTGAATTAACACGAA-----TCTGTAGCAAAAGGATCAACTAGC 706

RESULT 29

US-09-252-088-13

; Sequence 13, Application US/09252088

; Publication No. US20030031682A1

GENERAL INFORMATION:  
APPLICANT: BRODEUR, Bernard R.  
APPLICANT: RIOUX, Clément  
APPLICANT: BOYER, Martine  
APPLICANT: CHARLEBOIS, Isabelle  
APPLICANT: HAMEL, José  
APPLICANT: MARTIN, Denis  
TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS  
FILE REFERENCE: 8331-9002  
CURRENT APPLICATION NUMBER: US/09/252,088  
EARLIER FILING DATE: 1999-02-18  
EARLIER APPLICATION NUMBER: US/60/075,425  
EARLIER FILING DATE: 1998-02-20  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 13  
LENGTH: 5215  
TYPE: DNA  
ORGANISM: group B streptococcus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (3)..(122)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (133)..(2511)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (367)..(2511)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: Complement((2716)..(2946))  
FEATURE:  
NAME/KEY: CDS  
LOCATION: Complement((22995)..(3252))  
FEATURE:  
NAME/KEY: CDS  
LOCATION: Complement((3299)..(3676))  
FEATURE:  
NAME/KEY: CDS  
LOCATION: Complement((3837)..(4124))  
FEATURE:  
NAME/KEY: CDS  
LOCATION: Complement((4351)..(5214))  
US-09-252-088-13  
Alignment Scores:  
Pred. No.: 2,96e-82 Length: 5215  
Score: 1023.50 Matches: 289  
Percent Similarity: 46.05% Conservative: 107  
Best Local Similarity: 33.60% Mismatches: 253  
Query Match: 25,47% Indels: 212  
DB: 10 Gaps: 33  
US-09-765-272A-66 (1-763) x US-09-252-088-13 (1-5215)  
QY 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21  
Db 109 AGTTACCACTTGGTAAAGCATCATATGGTCTAGCAACAAAGGAC---AATCAGATTGCC 165  
QY 22 TyrIleAspGlyAspGlnAlaGlyGlnLysAla-----GluAsnLeuThrProAspGlu 39  
Db 166 TATATTGATGACGCAAGGTAAGCAAGGCCCTTAAACAAACAAACAAACGATGGATCAA 225  
QY 40 ValSerLysArgGluGlyIleAsnAlaGluGln\*\*\*ValIleLysIleThrAspGlnGly 59  
Db 226 ATCAGTCTGAAGAAGGATCTCTGCTGAACAGATCTGTCTGAAATTAATCTGACCAAGGC 285  
QY 60 TyrValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAla 79  
Db 286 TATGTGACCTCACAGGATGACCATTTATCATTTTACATGGGAAAGTTCTTATGATGCG 345  
QY 80 IleIleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIle 99  
Db 80 IleIleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIle 99

346 ATTATTAGTGAAGAGTGTGTTGATGACGGATCCTAATTACCGTTTTAAACAATCAGACGTT 405  
 100 ValAsnGluIleLysGlyTyrValIleLysValAsnGlyTyrValTyrLeu 119  
 406 ATCAATGAAATCTTAGACGGTTACGTTTAAAGTCAATGGCACTATTATGTTACCTC 465  
 120 LysAspAlaAlaHisAlaAspAsnIleArgThrLysGluIleLysArgGln 137  
 466 AAGCCAGGTAGTAAGCGCAAAACATTCGAACCAACAACAATTCCTCAGCAAGTAGCC 525  
 138 -----LysGlnGluArgSerHisAsn-----HisAsnSerArgAla 149  
 526 AAAGGACTAAAGAGCTAAAGAAAAGGTTTAGCTCAAGTGGCCCATCTCAGTAAAGAA 585  
 150 Asp---AsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrAspAspGlyTyr 168  
 586 GAAGTTTCGGCAGTCAATGAAGCAAAAGACAAAGACGCTATCTACAGCATGGCTAT 645  
 169 IlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleValProHisGly 188  
 646 ATTITTAGTCCGACAGATATCATTTGATGATTTAGGAGATGCTTATTAGTACCTCATG 705  
 189 AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGlu 208  
 706 AATCATTATCATTTATCTTAAAGATTTGTCTCCAGTGAGCTAGCTGCTGCACAA 765  
 209 AlaTyrTrpAsnGlyLysGln-----GlySerArgProSerSerSerSerTyrAsn 226  
 766 GCCTACTGGAGTCAAAACAAGGTCGAGGTCTAGACGGTCTGATTACCGCCGACCA 825  
 227 Ala-----AsnProAlaGlnProArg 233  
 826 GCCCCAGGTGTAGAAAGCCCAATTCCTGATGTGACGCCCTAACCTGGACAAGTCA 885  
 234 LeuSerGluAsnHisAsnLeuThrValThrPro-----Thr 245  
 886 CAGCCAGATACGGTGGCTATCATCCAGCGCTCTAGCGCCAAATGATGCGTCAAAAAC 945  
 246 TyrHisGlnAsn-----GlnGlyLysAsnIleSerSerLeuLeuArgGluLeuTyr 262  
 946 AAACACCAAGACATGAGTTTAAAGGAAAACCTTTAAGGAACCTTTAGATCAACTAC 1005  
 263 AlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGln 282  
 1006 CGCTTGTGATTGAATACCTCTGTGGAAAGATGGTGTGATTTTGAACCACTAC 1065  
 283 IleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIle 302  
 1066 GTGATCAAAATCAACCGCTTTTGGGTATGTGGTGCCTCATGGAGATCATTTATATTAT 1125  
 303 ProTyrGluGlnMetSerGluLeuGluLysArgIleAla---ArgIleIleProLeuArg 321  
 1126 CCAAGAGTCAGTTATCACTCTTGAATGGAAATTAGCATCGATCACTTAGCTGGCCAA 1185  
 322 TyrArgSerAsnHisTrpValProAspSerArgProGlu-----GlnProSerProGln 339  
 1186 ACTGAGGACAAT-----GACTCAGGTCAGAGCACTCAAAACCATCAGATAA 1233  
 340 SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer 359  
 1234 GAAGTGACACATACCTTTCTTGTGTCATCGCATCAAAAGCTTACGGAAAAGGCTTAGAT 1293  
 360 AsnProIleAspGluLys-----LeuValLysGluAlaValArgLysVal--- 374  
 1294 AAACCATATGATACGAGTATGCTTATGTTTAAAGAAATCCATTCATTTCAGTGGAT 1353  
 375 -----GlyAspGlyTyr----- 378  
 1354 AAATCAGGAGTTACAGCTAAACACGAGATCATTTCCACTATATAGATTGGAGAACTT 1413  
 379 ---ValPheGluGluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGlu 397  
 1414 GAACAATATGATGGTGGTTCGTAACCTGGTGAAGCAAAAGGTCAAGCTGATGAG 1473

QY 398 ThrAlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeu-----SerHisLys 415  
 Db 1474 CTTGCTGCTGCTTTGGATCAGGAACAGGCAAGAAAAACACCTCTTTGACATAAAAA 1533  
 QY 416 LeuGlyAlaLysLysThrAsp-----LeuProSerSerAspArg 428  
 Db 1534 GTGAGTCGCAAGTAACAAGATGTAAGTGGCTATATGATGCCAAAAAGATGTTAAG 1593  
 QY 429 GluPheTyr-----AsnLysAlaTyrAspLeu-----LeuAlaArgIleHisGlnAsp 444  
 Db 1594 GACTATTTCTGCTGCTGATCAACTTGAATTGACTCAGATTGCTTTGCCGCAACAAGAA 1653  
 QY 445 LeuLeuAspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGlu--- 463  
 Db 1654 CTAATGCTTAAAGATAAGAACGATTACCGTTATGACATTTGTGACACACAGGTATTGAGCCA 1713  
 QY 464 ArgLeuLys---AspVal\*\*Ser----- 470  
 Db 1714 CGACTTGTGTAGATGTGTCAAGTCTGCCGATCGATGCTGTAATGCTACTTACGATACT 1773  
 QY 471 -----AspLysValLysLeuVal----- 476  
 Db 1774 GGAAGTTCTGTTTATCCACATATTTGATCATATCCATCTCGTTCCGTATTCATGTTG 1833  
 QY 477 -----\*\*\*AspIleLeuAlaPheLeuAlaProIleArgHisProGlu----- 490  
 Db 1834 ACGCGGATCAGATTGCAACAGTCAAGTATGTATGATCAACACCCCGAAGTTCTGCTCGGAT 1893  
 QY 490 ----- 490  
 Db 1894 GTATGCTTAAGCCAGGCGCATGAAGTCAAGTTCGTCGTCTATCCAAATGTTACGCTCTT 1953  
 QY 491 -----ArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIleGlnVal 508  
 Db 1954 GATAAACGTCGTGTTATGCCAACTGCAAAATTTATCCATCTCTGCAAGAGTTCAAAA 2013  
 QY 509 AlaLysLeuAlaGlyLysTyrThrGluAspGlyTyrIlePheAspProArgAspIle 528  
 Db 2014 GCCCTAGCAGAGGTCGTTTTCGCAACACGACGCGCTATATTTTCGATCCACGAGATGT 2073  
 QY 529 ThrSerAspGlu-----GlyAspAlaTyrValThrProHisMetThrHis 543  
 Db 2074 TTGCCAAAGAAACTTTTGTATGGAAGATGCTCTTTAGCATCCCAAGACGATGGC 2133  
 QY 544 Ser-----HisTrpIleLysLysAspSerLysSerGluAlaGluArgAlaAlaGln 561  
 Db 2134 AGTTCATTGAGAACCATTAATAAATCTGATCTATCCCAAGCTGAGTGGCAACAGTCAA 2193  
 QY 562 AlaTyrAlaLysGluLysGlyLeu-----ThrProProSerThr 574  
 Db 2194 GAGTTATTGGCAAGAAAATACTGCTGATGCTACTGATACGGATAAACCAGAAAGAAAG 2253  
 QY 575 AspHisGlnAspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgVal 594  
 Db 2254 CAACAGCAGATAGACGATGAACCAACCAAGCAAGTGAAGCC-----AGTAAAGAA 2307  
 QY 595 LysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGlu 614  
 Db 2308 GAAAAAGAAATCAGATGACTTTATAGACAGTTTACCA----- 2343  
 QY 615 ValLysAsnGlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGlu 634  
 Db 2343 ----- 2343  
 QY 635 TrpPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuLeuAla 654  
 Db 2344 -----CACTATGCTCTAGATAGACCA-----ACCTAGAAAGATCATATCAAT 2385  
 QY 655 ThrVal-----LysTyrTyrValGluHisProAsnGluArg 666  
 Db 2386 CAATTAGCAAAAAGCTAATATATCGATCTTAAGTATCTCTATTTCCTCAACCA----- 2436

QY 667 ProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGln---ArgAsnLysAsn 695  
 Db 2437 -----GAAGGTGTCCTCAATTTTATATAATAAAAT 2463

QY 686 GlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLys 705  
 Db 2464 GGTGAATTGGTAACCTATCATATAT-CAGACACTTCACAAATAAACCCCTAACCAAAAGA 2522

RESULT 30  
 US-10-340-792-13  
 ; Sequence 13, Application US/10340792  
 ; Publication No. US20030228323A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BRODEUR, Bernard R.  
 ; APPLICANT: RIOUX, Clement  
 ; APPLICANT: BOVER, Martine  
 ; APPLICANT: CHARLEBOIS, Isabelle  
 ; APPLICANT: HAMEL, Josee  
 ; APPLICANT: MARTIN, Denis  
 ; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS  
 ; FILE REFERENCE: 8331-9002  
 ; CURRENT APPLICATION NUMBER: US/10/340,792  
 ; PRIOR FILING DATE: 2003-01-13  
 ; PRIOR APPLICATION NUMBER: US/09/252,088B  
 ; PRIOR FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US/60/075,425  
 ; PRIOR FILING DATE: 1998-02-20  
 ; NUMBER OF SEQ ID NOS: 45  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 13  
 ; LENGTH: 5215  
 ; TYPE: DNA  
 ; ORGANISM: group B streptococcus  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (3)..(122)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (133)..(2511)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (367)..(2511)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: Complement((2716)..(2946))  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: Complement((2295)..(3252))  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: Complement((3299)..(3676))  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: Complement((3837)..(4124))  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: Complement((4351)..(5214))  
 ; US-10-340-792-13

Alignment Scores:  
 Pred. No.: 2,96e-82 Length: 5215  
 Score: 1023.50 Matches: 289  
 Percent Similarity: 46.05% Conservative: 107  
 Best Local Similarity: 33.60% Mismatches: 253  
 Query Match: 25.47% Indels: 212  
 DB: 16 Gaps: 33

US-09-765-272a-66 (1-763) x US-10-340-792-13 (1-5215)

QY 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21  
 Db 109 AGTTACCACTTGTGAAGCATCATATGGGTCTAGCAACAAAGGAC---AATCAGATTGCC 165

QY 22 TyrIleAspGlyAspGlnAlaGlyGlnLysAla-----GluAsnLeuThrProAspGlu 39  
 Db 166 TATATTGATGACGCAAGAGGTAGGCAAAAGCCCTAAACAAAACAAACGATGATCAA 225

QY 40 ValSerLysArgGluGlyIleAsnAlaGluGln\*\*ValIleLysIleThrAspGlnGly 59  
 Db 226 ATCAGTGTCTGAAGAGGCATCTCTGCTGAACAGATCGTAGTCAAAATTTACTGACCAAGGC 285

QY 60 TyrValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProThrAspAla 79  
 Db 286 TATGTGACCTCACAGGTGACCATATATCTTTTACAAATGGGAAAGTTCTTATGATGCG 345

QY 80 IleIleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLysAspSerAspIle 99  
 Db 346 ATTATTAGTGAAGAGTTGTTGATGACGATCTTAATTTACCGTTTAAACAATCAGACGT 405

QY 100 ValAsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrValTyrIleu 119  
 Db 406 ATCAATGAATCTTAGACGGTTAGCTTATTAAAGTCAATGGCAACTATTATGTTACCTC 465

QY 120 LysAspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGln----- 137  
 Db 466 AAGCCAGGTAGTAGGCGCAAAAACATTCGAACCAACAAATTTGCTGACCAAGTAGCC 525

QY 138 -----LysGlnGluArgSerHisAsn-----HisAsnSerArgAla 149  
 Db 526 AAAGGAACCTAAAGAAAGCTAAAGAAAGGTTTAGCTCAAGTGGCCCATCTCAGTAAAGAA 585

QY 150 Asp---AsnAlaValAlaAlaArgAlaGlnGlyArgTyrThrAspAspGlyTyr 168  
 Db 586 GAAGTTGCGCGCAGTCAATGAAGCAAAAGCAAGGACGCTATCTACAGACGATGGCTAT 645

QY 169 IlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleValProHisGly 188  
 Db 646 ATTTTGTAGTCCGACATATCATGATGATGATGATGATGATGATGATGATGATGATGAT 705

QY 189 AspHisTyrHisTyrIleProLysAsnGluLeuSerAlaSerGluLeuAlaAlaGlu 208  
 Db 706 AATCACTATCATATATCTCTAAAGAGGTTGTTCTCAAGTGGCTGCTGCTGCTGCTGCT 765

QY 209 AlaTyrTyrAsnGlyLysGln-----GlySerArgProSerSerSerSerSerTyrAsn 226  
 Db 766 GCCTACTGGAGTCAAAACAAAGGTCGAGGTGCTAGACCGTCTGATTACCGCCGACACCA 825

QY 227 Ala-----AsnProAlaGlnProArg 233  
 Db 826 GCCCAGGTGCTGAGAAAGCCCAATTCCTGATGTGAGCGCTAACCCCTGGACACGTCAT 885

QY 234 LeuSerGluAsnHisAsnLeuThrValThrPro-----Thr 245  
 Db 886 CAGCCAGATAACGGTGGCTATCATCCAGCGCTCTTAGGCCAAATGATGGTGCACAAAC 945

QY 246 TyrHisGlnAsn-----GlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyr 262  
 Db 946 AAACACCAAGAGATGAGTTTAAAGGAAAAACCTTTAAGGAACCTTTTAGTCAACTACAC 1005

QY 263 AlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGln 282  
 Db 1006 CGTCTTGATTGAAATACCGTCATGGGAAGAGATGGGTGTTGATTTTGAACCGACTCAA 1065

QY 283 IleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIle 302  
 Db 1066 GTGATCAAAATCAACGCTTTTGGGTATGTGGTCTCATGGAGATCATATATCATATATATC 1125

QY 303 ProTyrGluGlnMetSerGluLeuGluLysArgIleAla---ArgIleIleProLeuArg 321  
 Db 1126 CCAGAAGTCACTTATCATCTCTGAAATGGAATAGCAGATCGATCTAGCTGGGCCAA 1185

QY 322 TyrArgSerAsnHisTyrValProAspSerArgProGlu-----GlnProSerProGln 339  
 Db 1186 ACTGAGGCAAT-----GACTCAGGTTTCAGACACTCAAAACCATCAGATAAA 1233

QY 340 SerThrProGluProSerProSerProGlnProAlaProAsnProGlnProAlaProSer 359





80 IleIleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIle 99  
301 ATTATTAGTGAAGAGTGTGATGACGGATCTTAATACCAATTTTAAACAATACACAGCT 360  
100 ValAsnGluIleLysGlyGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeu 119  
361 ATCAATGAACTTAGACGGTTACGTTATTAAAGTCATGGAACATATTATGTTTACCTC 420  
120 LysAspAlaAlaHisAlaAspAsnIleArgThrLysGluGluIleLysArgGln----- 137  
421 AAGCCAGGTAGTAAGCGCAAAACATTCGAACCAACAAACAAATTTGCTGAGCAAGTAGCC 480  
138 -----LysGlnGluArgSerHisAsn-----HisAsnSerArgAla 149  
481 AAGGAACATAAAGAAAGTAAAGAAAAAGGTTTAGCTCAAGTGGCCCATCTCAGTAAAGAA 540  
150 Asp---AsnAlaValAlaAlaArgAlaGlnGlyArgTyrThrThrAspAspGlyTyr 168  
541 GAAGTTGGCGGAGTCATGAAGCAAAAGACCAAGGACCTATACACAGCATGGCTAT 600  
169 IlePheAsnAlaSerAspIleIleGluAspThrGlyAspAlaTyrIleValProHisGly 188  
601 ATTATTTAGTCGACAGATATCATTTGATGATTTAGGAGATGCTTATTAGTACCTCATGTT 660  
189 AspHisTyrHisTyrIleProLysAsnGlnLeuSerAlaSerGluLeuAlaAlaGlu 208  
661 AATCATCATATTATATCTTAAAGAAATTTGCTCAAGTAGAGTAGCTAGCTGCACAA 720  
209 AlaTyrTrpAsnGlyLysGln-----GlySerArgProSerSerSerSerTyrAsn 226  
721 GCCTACTGGAGTCAAAACAAAGGTCGAGGTGCTAGACCGTCTGATTACCGCCGACACA 780  
227 Ala-----AsnProAlaGlnProArg 233  
781 GCCCCAGGTCGTAGAAAGCCCAATTCCTGATGTGACGCCCTAACCTGGCAAGGTAT 840  
234 LeuSerGluAsnHisAsnLeuThrValThrPro-----Thr 245  
841 CAGCCAGATACCGTGGTTATCATCCAGCGCTCTAGCGCAATGATGCGTCACAAAC 900  
246 TyrHisGlnAsn-----GlnGlyGluAsnIleSerSerLeuLeuArgGluLeuTyr 262  
901 AAACACAAAGAGATGATTTAAAGGAAAAACCTTTAAGAACTTTTATAGATCATCTACAC 960  
263 AlalysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAspProAlaGln 282  
961 CGCTTTCATTGAAATACCGTCATGTGGAAGACATGGTTGATTTTGAACCGACTCAA 1020  
283 IleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIle 302  
1021 GTGATCAAAATCAAAACGCTTTTGGGTATGTGGTCTCATGGAGATCATATCATATTATC 1080  
303 ProTyrGluGlnMetSerGluLeuGluLysArgIleAlaArgIleIleProLeuArgTyr 322  
1081 CCAAGAGTCAGTTATCACTCTTGAATGGAATAGACAGAT-----CGATAC 1128  
323 ArgSerAsnHisTrpValProAspSerArgProGluGlnProSerProGlnSerThrPro 342  
1129 TTAGCCGCGC-----CAAAGTATGATC 1149  
343 GluProSerProSerProGlnProAlaProAsnProGlnProAla----- 357  
1150 AAGCACTCAGGTTTCAGATCACTCAAAACCATCAGATAAAGAAAGTGACACATACCTTTCTT 1209  
358 -----ProSerAsnProIleAspGluLys--- 365  
1210 GGTATCGCATCAAGCTTACGAAAAAGGCTTAGATGGTAAACCATATGATACAGAGTAT 1269  
366 -----LeuValLysGluAlaValArgLysVal----- 374  
1270 GCTATGTTTTTAGTAAGAATCCATTCATTCAGTGGATAAATCAGGAGTTTACAGCTAAA 1329  
375 ---GlyAspGlyTyr-----ValPheGluGluAsnGly 384

1330 CACGGAGATCATTTCCACTATATAGGATTTGGAGAACTTGAACAATATAGATTGGATGAG 1389  
385 ValSerArgTyrIleProAlaLysAspLeuSerAlaGlnThrAlaAlaGlyIleAspSer 404  
1390 GTTCGCTAACTGGGTGAAGCAAAAGGTCAAGCTGATGAGCTGTTGCTGCTGGATCAG 1449  
405 LysLeuAlaLysGlnGluSerLeu-----SerHisLysLeuGlyAlaLysLysThrAsp 422  
1450 GAACAGGCAAGAAAAACCACTCTTTGACACTAAAAAAGTGAAGTCCCAAGTAAACAAA 1509  
423 -----LeuProSerSerAspArgGluPheTyrAsnLysAlaTyr 435  
1510 GATGTAAAGTGGGCTATATTATGCAAAAGATGGCAAGGACTATTCTATCTGCTCGTTAT 1569  
436 AspLeu-----LeuAlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArg 451  
1570 CAACITGATTGACTCAGATTGCTTTGGCAACAAAGAACTAATGCTTAAAGATAAAGAA 1629  
452 GlnValAspPheGluAlaLeuAspAsnLeuGlu-----ArgLeuLys---AspVal\*\*\* 469  
1630 CATTACCGTTATGACATTTGTTGATCAGGCATTGAGCCAGCACTTGTCTGTAGATGTCA 1689  
470 Ser----- 470  
1690 AGTCTGCCGATGCTGCTGTAATGCTACTTACGATACTGGAAGTTGCTTTGTTTATCCCA 1749  
471 -----AspLysValLysLeuVal-----\*\*\*AspIleLeuAla 481  
1750 CATATTGATCATCATTCATGCTGCTTCGTTATTCATGTTGACGCAATCAGATTGCAACA 1809  
482 PheLeuAlaProIleArgHisProGlu----- 490  
1810 ATCAAGTATGTGATGCAACACCCCGAAGTTCGTCCGAGTGTATGTTTAAGCCAGGCGAT 1869  
491 -----ArgLeuGlyLysPro 495  
1870 GAAGAGTCAGGTTGCTGCTCATTCAAATGTTAGCCCTCTTGATAAAGCTGCTGGTATGCCA 1929  
496 AsnAlaGlnIleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyr 515  
1930 AACTGGCAATTTATCCATTTCTGCTGAAGATTTCAAAAAGCCCTAGCAGAGGTCGCTTT 1989  
516 ThrThrGluAspGlyTyrIlePheAspProArgAspIleThrSerAspGlu----- 532  
1990 GCAGCACAGACGGCTATATTTTCGATCCAGAGATGTTTGGCAAAAGAACTTTTGTGA 2049  
533 -----GlyAspAlaTyrValThrProHisMetThrHisSer-----HisTrpIleLys 548  
2050 TGAAGAGATGGCTCCTTTAGCATCCCAAGCAGATGGCAGTTTCATTGAGAACCATTAA 2109  
549 LysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAlaLysGluLysGly 568  
2110 AAATCCGATCTATCCCAAGCTGAGTGGCAACAAAGCTCAAGATTATTGGCAAAAGAAAT 2169  
569 Leu-----ThrProProSerThrAspHisMetThrHisSer-----HisTrpIleLys 581  
2170 GCTGTGTGCTACTGATACGGATAAACCCTGAAGAAAGCAACAGCAGCATAGAGCAAT 2229  
582 ThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaLysLysValPro 601  
2230 GAAAAACCAACAGCCCAAGTAGAGCC-----AGTAAAGAAAGAAAGCAATCAGATGACTTT 2283  
602 LeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIle 621  
2284 ATAGACAGTTTACCA----- 2298  
622 IleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeuTyr 641  
2299 -----GACTATGCTGTAGAT 2313  
642 GluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrVal----- 656

```

2314 AGAGCA-----ACCTAGAGATCATCATCAATAGACACAAAAGCTAAT 2361
657 -----LysTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPhe 673
2362 ATCGATCCTAAGTATCTCATTTTCCAACCA-----2391
674 GlyAsnAlaSerAspHisValGln--ArgAsnIysAsnGlyGlnAlaAspThr 690
2392 -----GAAGGTGTGCCAATTTTATAATAAAAAATGTTGAATGGTAACT 2433

RESULT 32
US-09-884-465A-9
; Sequence 9, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: Patent In version 3.1

```

Qy	141	ArgSerHisAsnHisAsnSerArgAlaAspAsnAlaValaAlaAlaArgAlaGlnGly	160
Db	418	CATGTCAAAGAT---AATGGAAGGTTAACTCTAATGTTGCTGTAGCAAGGCTTCAGGA	474
Qy	161	ArgTyrThrThrAspAspGlyTyrIlePheAsnAlaSerAspIleIleGluAspThrGly	180
Db	475	CGATATACGCAAAATGATGGTTATGCTTTAATCCAGCTGATATATTCGAAGATACGGGT	534
Qy	181	AspAlaTyrIleValProHisGlyAspHisTyrHisTyrIleProLysAsnGluLeuSer	200
Db	535	AATGCTTATATCGTTCCTCATCGAGGTCACATACCTACATCCCAAAAGCGATTATCT	594
Qy	201	AlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLysGlnGlySerArgProSer	220
Db	595	GCTAGTGAATTAGCAGCAGCTAAAGCACATCTGGCTGGAAA---AATATGCAACCG	648
Qy	221	SerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSerGluAsnHisAsnLeu	240
Db	649	AGTCAGTTAGCTATCTTCAACACCTTCCTCCA---TCCTCTT	687
Qy	241	ThrValThrPro---ThrTyrHisGlnAsnGlnGlyGluAsnIleSerSerLeuLeuArg	259
Db	688	CCAATCAATCCAGGAATTCACAT-----	711
Qy	260	GluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeuIlePheAsp	279
Db	712	-----GAGAAACAT---GAAGAAGATGATACGGATTTGAT	744
Qy	280	ProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGlyAsnHisTyr	299
Db	745	GCTAATCGTATATTCGCTGAAGATGAATCAGGTTTTCATGAGTCACGAGACCAAT	804
Qy	300	HisPhe-----IleProTyrGluGlnMetSerGluLeuGluLysArgIle	314
Db	805	CATTATTCTTCAGAAGGACTTCACAGAGAGCAAAATTAAGGCTGGCGAAACATTTA	864
Qy	315	AlaArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAsp-----SerArg	332
Db	865	GAGGAAGTT-----AAAATAGTCATAATGGATAGATTTCTTTGTCTATCT	909
Qy	333	ProGluGlnProSerProGlnSerThrProGluProSerProSerProGlnProAla---	351
Db	910	CATGAACAGGATTATCCACGTAATGCCAAAGAAATGAAGATTTAGAATAAAAAATCGAA	969
Qy	351	-----	351
Db	970	GAATAAAATTGCTGGCATATTGAACAATATGGTGTCAAACGTGAAAGTATGTCTGCAAT	1029
Qy	352	-----ProAsnProGlnProAlaProSerAsnProIleAsp	363
Db	1030	AAAGAAAAAATGCGATTATTTATCCGCATGGAGATCACCATCATGCAGATCCGATTGAT	1089
Qy	364	Glu-----	364
Db	1090	GAACATAAACCGGTTGGAATTGGTCATCTTCACAGTAACATGAACGTGTTTAAACCGAA	1149
Qy	364	-----	364
Db	1150	GAAGGAGTTGCTAAAAAAGAGGGAATAAGTTTATCTCGAGAGAAGATTAACGAATGTT	1209
Qy	365	----LysLeuValLys-----GluAlaValArgLysValGly-----	368
Db	1210	GTTAATTTGTTAAAAAATAGTAGCTTTTAATATCAAAACCTTTACTCTGCCAATGGTCAA	1269
Qy	369	-----	375
Db	1270	AAACGCGTTTCTTTTAGTTTTCGCTCGAATTGGAGAAAAAATTAGTATCAATATGCTA	1329
Qy	376	-----AspGlyTyrValPheGluGlu-----	382
Db	1330	GTAAAAATTAAACACGAGATGGAAGAAGTATTGGAGAAAGATATCTCGTAAGATTTTGA	1389
Qy	383	AsnGlyValSerArg-----TyrIleProAlaLysAsp	393

```
Db 1390 GAAGGAGTAGGAATATTCGAACTTTGATTAGATCAACTTATTATTCAGGACAAACA 1449
Qy 394 LeuSerAlaGluThrAlaAla-----GlyLeuAspSerLysLeuAla 407
Db 1450 TTAAAGTATATCTATCGCTTCAGAAAGATTATCCAGAGTAGTATTGATGATACATTTACA 1509
Qy 408 LysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSerSerAsp 427
Db 1510 GTTCCAACTCTTTAGCTTACAAATGGCCAGTCAAAACAGATTCTCTCTCCATGCA 1569
Qy 428 ArgGluPheTyr-----AsnLysAlaTyrAspLeuLeu 438
Db 1570 GGGGATACTATTATAGAGTGAACCTCAATTTGCGAGTGCTTAAGGAAGTATGATGCTTTA 1629
Qy 439 AlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPheGluAla-- 457
Db 1630 GTCAGAGTCTTTGATGAATTCATGCAATGCTTATTAGAAAATAACTATATAAGTTGGT 1689
Qy 458 -----LeuAspAsnLeuLeuGluArgLeuLysAspVal***SerAspLys 472
Db 1690 GAAATCAAAATACCGATTCCGAAATTTAAACCAAGGAACRACCAAGACGGCCGGAATAAA 1749
Qy 473 ValLysLeuVal***AspIleLeuAlaPheLeu----- 483
Db 1750 ATTCTGTAACTTCAATGCGAAATGCTTATTGGACAATCAATCGACTTATATTTGTGAA 1809
Qy 484 AlaProIleArgHisProGlu---ArgLeuGlyLysProAsnAlaGlnIleThrTyrThr 502
Db 1810 GTACCTATCTTGAAAAAGAAATCAAACTGATAACCAAGTATTCTACCACTTTAAA 1869
Qy 503 AspAspGluIleGlnVal---AlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyr 521
Db 1870 AGGAATAAAGCACAGAAAACTCAAACTT-----GATGAAAAG 1908
Qy 522 IlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMet 541
Db 1909 GTAGAAAGAACCAAG-----ACTAGTGAGAG- 1935
Qy 542 ThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGln 561
Db 1936 -----GTAGAAAAGAAAGAACTTCTGAAACTGGGAATAGTACTAGTAAT 1980
Qy 562 AlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsn 581
Db 1981 TCAACGTTAGAGAAAGTTCTCAGTGGATCCTGTACAAGAA----- 2022
Qy 582 ThrGluAlaLysGlyValGluAlaIleTyrAsnArgValLysAlaAlaLysLysValPro 601
Db 2023 AAAGTAGCAAAATTTGCTGAAAGTTAT-----GGGATGAAG 2058
Qy 602 LeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIle 621
Db 2059 CTAGAAAATGTCTTGTTAATATGGACGAAACAAATTGA----- 2097
Qy 622 IleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAspGluGlyLeuTyr 641
Db 2098 -----TTATAT 2103
Qy 642 GluAlaProLysGlyTyrThrLeuGluAspLeuLeuAlaThrValLysTyrTyrValGlu 661
Db 2104 ---TTACCATCGGAGAGAGTCAATTA-----AAGAAATATGGCAGAT 2142
Qy 662 HisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGln 681
Db 2143 TTTACAGAGAGACCTCAA-----GGAATGGTGAAAT-----AAA 2181
Qy 682 ArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGlu---LysProSerGluGluLys 700
Db 2182 CCATCTGAAATGGAAGAAAGTATCTACTGGAACAGTTGAAACCAACCAACAGAAATAAA 2241
Qy 701 ProGlnThrGluLysProGluGluThrProArgGluGluLysProGlnSerGlu--- 719
```

```
Db 2242 CCAGCAGATTCTTTACCAGAGGCCAACCAACGAAAAACCTGTAAAAACCAAGAACTCAACG 2301
Qy 720 -----LysProGluSerProLysProThrGluGluProGluGluSerPro 734
Db 2302 GATAATGGAATGTTGAATCCAGAGGGGAATGTGGGAGTGACCTTATGTTAGATTGCGCA 2361
Qy 735 GluGluSerGluGluProGlnValGluThrGluLysValGluGluLysLeuArgGlu 753
Db 2362 TTAGAG---GAAGCTCCAGCAGTA-----GATCCTGTACAGAAAAAATTAGAAAAA 2409
RESULT 33
US-09-769-736-23
; Sequence 23, Application US/09769736
; Publication No. US20030138775A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21089wo
; CURRENT APPLICATION NUMBER: US/09/769,736
; PRIOR FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: GB 9816335.5
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125163
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Streptococcus agalactiae
US-09-769-736-23
Alignment Scores:
Pred. No.: 7,09e-70 Length: 1146
Score: 877.00 Matches: 185
Percent Similarity: 62.95% Conservative: 41
Best Local Similarity: 51.53% Mismatches: 81
Query Match: 11.82% Indels: 52
DB: 10 Gaps: 11
US-09-765-272A-66 (1-763) x US-09-769-736-23 (1-1146)
Qy 2 SerTyrGluLeuGlyArgHisGlnAlaGlyGlnValLysLysGluSerAsnArgValSer 21
Db 64 AGTTACCAGCTTGGTAAGCATCATATGGTCTAGCACAAAGGAC---AATCAGATTGCC 120
Qy 22 TyrIleAspGlyAspGlnAlaGlyGlnLysAla-----GluAsnLeuThrProAspGlu 39
Db 121 TATATTGATGATACCAAGGTAAGTAAAGCCCTTAAACAAACAAACGATGATCAA 180
Qy 40 ValSerLysArgGluGlyIleAsnAlaGluGln***ValIleLysIleThrAspGlnGly 59
Db 181 ATCAGTCTCAAGAGGCACTCTCTGCTGACAGATCGTAGTCAAAATTACTGACCAAGGT 240
Qy 60 TyrValThrSerHisGlyAspHisTyrHisTyrTyrAsnGlyLysValProTyrAspAla 79
Db 241 TATGTTTACCTCACACGGTGACCATTATCATTTTACAATGGGAAAGTCTCTTATGATGCG 300
Qy 80 IleIleSerGluGluLeuLeuMetLysAspProAsnTyrGlnLeuLysAspSerAspIle 99
Db 301 ATTATTAGTAAGAGTGTGTGATGACGGATCCTTAATTTACCATTTTAAACAATCAGACGTT 360
Qy 100 ValAsnGluIleLysGlyTyrValIleLysValAsnGlyLysTyrTyrValTyrLeu 119
Db 361 ATCAATGAATCTTAGACGGTTAGCTTATTAACTCAATGGCAACTATTATGTTTACCTC 420
Qy 120 LysAspAlaAlaHisAlaAspAsnIleArgThrLysGluGluLysLysArgGln----- 137
Db 421 AAGCCAGGTAGTAAGGCGCAAAAACATTCGAACCAACAAACAAACAAATTCGTGAGCAAGTAGCC 480
```



; CURRENT APPLICATION NUMBER: US/10/324,143  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: 60/341,252  
; PRIOR FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 160  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 819  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: modified Streptococcus pneumoniae nucleotide sequence  
US-10-324-143-10

Alignment Scores:  
Pred. No.: 1,376-65 Length: 819  
Score: 827.50 Matches: 169  
Percent Similarity: 72.83% Conservatives: 16  
Best Local Similarity: 66.54% Mismatches: 28  
Query Match: 20.59% Indels: 41  
DB: 16 Gaps: 5

US-09-765-272A-66 (1-763) x US-10-324-143-10 (1-819)

QY 498 GlnIleThrTyThrAspAspGluLeuValAlaLysLeuAlaGlyLysTyThrThr 517  
Db 4 CAATATTACCTACCTGATGAGATTCAGGTAGCCAGTTGGCAGGCAAGTACACAA 63  
QY 518 GluAspGlyTyThrPheAspProArgAspIleThrSerAspGluGlyAspAlaTyVal 537  
Db 64 GAAGACGGTTATATCTTTGAT 84  
QY 538 ThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArg 557  
Db 85 -----ACTAGTTGGATTAAAGAGATAGTTTGTCTGAAGCTGAGAGA 126  
QY 558 AlaAlaAlaGlnAlaTyThrAlaLysGluLysGlyLeuThrProProSerThrAspHisGln 577  
Db 127 GCGGCAGCCAGGCTTATGCTAAGAGAAAGGTTTGAACCCCTCTTCGACACCCAC 186  
QY 578 AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaLysLeuValAlaLysValAla 597  
Db 187 GATTTCAGGAATACTAGGCAAAAGGAGCAGAGCTATCTAACCCGGGTGAAGCAGCT 246  
QY 598 LysLysValProLeuAspArgMetProTyThrAsnLeuGlnTyThrValGluValLysAsn 617  
Db 247 AAGAGGTGCACCTTGATCGTATGCTTACATCTTCAGTATCTGTAGAGTCAAAAC 306  
QY 618 GlySerLeuIleIleProHisTyThrAspHisTyThrHisAsnIleLysPheGluTrpPheAsp 637  
Db 307 GGTAGTTTAAATCATACCTCATCATCATTACCAATCAATCAATCAATCAATCAATCA 366  
QY 638 GluGlyLeuTyThrGluAlaProLysGlyTyThrLeuGluAspLeuLeuAlaThrValLys 657  
Db 367 GAAGCCCTTTATGAGCACCTTAAGGGGTATAGTCTTGGAGATCTTTTGGCGACTCTCA 426  
QY 658 TyThrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 677  
Db 427 TACTATGTGAAA-----CCGGCG-----AACGCTAGT 453  
QY 678 AspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSer 697  
Db 454 GACCATGTTCTGAAAATAAG-----GCAGACCAAGATAGTAAACCTGAT 498  
QY 698 GluGluLysProGlnThrGluLysProGluGluThrProArgGluGluLysProGln 717  
Db 499 GAAGATAAGCAACATGATGAAGTAAGTGAAGCAACTCACCCCTGAATCTGATGAAAAGAG 558  
QY 718 SerGluLysProGluSerPro-----LysProThrGluGluProGlu 731  
Db 559 AATCACCGCTGGTTTAAATCTTCAGCAGATAATCTTTATAAACAAGCACTGATACGGA 618

QY 732 GluSerProGluGluSerGluGluProGlnValGluThrGlu 745  
Db 619 GAGACAGGAGGAAGGCTGAAGATACCACAGATGAGGCTGAA 660

RESULT 36  
US-10-324-143-6  
; Sequence 6, Application US/10324143  
; Publication No. US20030232976A1  
; GENERAL INFORMATION:  
; APPLICANT: HAMEL, JOSEPH  
; APPLICANT: CHARLAND, NATHALIE  
; APPLICANT: BRODEUR, BERNARD R.  
; APPLICANT: MARTIN, DENIS  
; APPLICANT: BLAIS, NORMAND  
; APPLICANT: OUELLETTE, CATHERINE  
; TITLE OF INVENTION: STREPTOCOCCUS ANTIGENS  
; FILE REFERENCE: 55190-56  
; CURRENT APPLICATION NUMBER: US/10/324,143  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: 60/341,252  
; PRIOR FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 160  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 2721  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: modified Streptococcus pneumoniae nucleotide  
; OTHER INFORMATION: sequence  
US-10-324-143-6

Alignment Scores:  
Pred. No.: 7,076-65 Length: 2721  
Score: 827.50 Matches: 168  
Percent Similarity: 72.44% Conservatives: 16  
Best Local Similarity: 66.14% Mismatches: 29  
Query Match: 20.59% Indels: 41  
DB: 16 Gaps: 4

US-09-765-272A-66 (1-763) x US-10-324-143-6 (1-2721)

QY 498 GlnIleThrTyThrAspAspGluLeuValAlaLysLeuAlaGlyLysTyThrThr 517  
Db 4 CAATATTACCTACCTGATGAGATTCAGGTAGCCAGTTGGCAGGCAAGTACACAA 63  
QY 518 GluAspGlyTyThrPheAspProArgAspIleThrSerAspGluGlyAspAlaTyVal 537  
Db 64 GAAGACGGTTATATCTTTGAT 84  
QY 538 ThrProHisMetThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArg 557  
Db 85 -----CGTGGTTGGATTAAAGAGATAGTTTGTCTGAAGCTGAGAGA 126  
QY 558 AlaAlaAlaGlnAlaTyThrAlaLysGluLysGlyLeuThrProProSerThrAspHisGln 577  
Db 127 GCGGCAGCCAGGCTTATGCTAAGAGAAAGGTTTGAACCCCTCTTCGACACCCAC 186  
QY 578 AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaLysLeuValAlaLysValAla 597  
Db 187 GATTTCAGGAATACTAGGCAAAAGGAGCAGAGCTATCTAACCCGGGTGAAGCAGCT 246  
QY 598 LysLysValProLeuAspArgMetProTyThrAsnLeuGlnTyThrValGluValLysAsn 617  
Db 247 AAGAGGTGCACCTTGATCGTATGCTTACATCTTCAGTATCTGTAGAGTCAAAAC 306  
QY 618 GlySerLeuIleIleProHisTyThrAspHisTyThrHisAsnIleLysPheGluTrpPheAsp 637  
Db 307 GGTAGTTTAAATCATACCTCATCATCATTACCAATCAATCAATCAATCAATCAATCA 366  
QY 638 GluGlyLeuTyThrGluAlaProLysGlyTyThrLeuGluAspLeuLeuAlaThrValLys 657

367 GAAGGCTTTATGAGCACCCTAAGGGTATAGTCTTGAGATCTTTGGGAGCTGTCAAG 426  
 658 TyrtValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 677  
 427 TACTATGTCGAAGGG-----GGGAACGCTAGT 453  
 678 AspHisValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 697  
 454 GACCATGTCGTAAGGG-----GCAGACCAAGATAGTAAACCTGAT 498  
 698 GluGluLysProGluThrGluLysProGluGluGluThrProArgGluGluLysProGlu 717  
 499 GAAGATAAGGAACATGATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 558  
 718 SerGluLysProGluSerPro-----LysProThrGluGluProGlu 731  
 559 AATCAGCTGGTTAAATCTTCAGCAGATATCTTTATAACCAAGCCTGATACGGAA 618  
 732 GluSerProGluGluSerGluGluProGluGluValGluThrGlu 745  
 619 GAGACAGAGGAAGAGCTGAAGATACCAAGATGAGGCTGAA 660

## RESULT 37

JS-10-324-143-4  
 ; Sequence 4, Application US/10324143  
 ; Publication No. US20030232976A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HAMEL, JOSEE  
 ; APPLICANT: CHARLAND, NATHALIE  
 ; APPLICANT: BRODEUR, BERNARD R.  
 ; APPLICANT: MARTIN, DENIS  
 ; APPLICANT: BLAIS, NORMAND  
 ; APPLICANT: OUELETTE, CATHERINE  
 ; TITLE OF INVENTION: STREPTOCOCCUS ANTIGENS  
 ; FILE REFERENCE: 55190-56  
 ; CURRENT APPLICATION NUMBER: US/10/324,143  
 ; CURRENT FILING DATE: 2002-12-20  
 ; PRIOR APPLICATION NUMBER: 60/341,252  
 ; PRIOR FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 160  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 816  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 ; OTHER INFORMATION: modified Streptococcus pneumoniae nucleotide  
 ; OTHER INFORMATION: sequence  
 JS-10-324-143-4

Alignment Scores:  
 Pred. No.: 3.88e-65 Length: 816  
 Score: 822.50 Matches: 168  
 Percent Similarity: 72.83% Mismatches: 17  
 Best Local Similarity: 66.14% Indels: 41  
 Query Match: 20.47% Gaps: 5  
 JB: 16

JS-09-765-272A-66 (1-763) x US-10-324-143-4 (1-816)  
 498 GlnIleThrTyrtValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 517  
 4 CAATTAACCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 63  
 518 GluAspGlyTyrtValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 537  
 64 GAAGACGGTTATATCTTTCAT----- 84

538 ThrProHisMetThrHisSerHisTrpIleLysAspSerLysSerGluAlaGluArg 557  
 85 -----ACTAGTTGATTAAGAAAGATGTTTGTCTGAGCTGAGAGA 126

558 AlaAlaAlaGlnAlaTyrtValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 577  
 127 GCGCAGCAGCCAGGCTTATGCTTAAAGAGAAAGGTTTGAACCCCTCTCTTCGACAGACACCAAG 186  
 578 AspSerGlyAsnThrGluAlaLysGlyValaGluAlaIleTyrtValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 597  
 187 GATTCAAGAAATATCTGAGGCAAAAGGAGCAAGCTATCTCAACCCGGGTGAAGCAGCT 246  
 598 LysLysValProLysAspArgMetProTyrtValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 617  
 247 AAGAGGTGCGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 306  
 618 GlySerLeuIleLeuProHisTyrtValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 637  
 307 GGTAGTTTAAATCATACCTCATCATGACCATTAACCAATCAATCAATCAATCAATCAATCAATCAAT 366  
 638 GluGluLysProGluThrGluAlaProLysGlyTyrtValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 657  
 367 GAAGGCTTTATGAGGCAACCTTAAGGGGTATAGTCTTTGAGGATCTTTTGGCGACTGTCAAG 426  
 658 TyrtValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 677  
 427 TACTATGTCGAA-----CCGCGG-----AACGCTAGT 453  
 678 AspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSer 697  
 454 GACCATGTCGTAAGGG-----GCAGACCAAGATAGTAAACCTGAT 498  
 698 GluGluLysProGluThrGluLysProGluGluGluThrProArgGluGluLysProGlu 717  
 499 GAAGATAAGGAACATGATGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 558  
 718 SerGluLysProGluSerPro-----LysProThrGluGluProGlu 731  
 559 AATCAGCTGGTTAAATCTTCAGCAGATATCTTTATAACCAAGCCTGATACGGAA 618  
 732 GluSerProGluGluSerGluGluProGluGluValGluThrGlu 745  
 619 GAGACAGAGGAAGAGCTGAAGATACCAAGATGAGGCTGAA 660

## RESULT 38

US-10-324-143-5  
 ; Sequence 5, Application US/10324143  
 ; Publication No. US20030232976A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HAMEL, JOSEE  
 ; APPLICANT: CHARLAND, NATHALIE  
 ; APPLICANT: BRODEUR, BERNARD R.  
 ; APPLICANT: MARTIN, DENIS  
 ; APPLICANT: BLAIS, NORMAND  
 ; APPLICANT: OUELETTE, CATHERINE  
 ; TITLE OF INVENTION: STREPTOCOCCUS ANTIGENS  
 ; FILE REFERENCE: 55190-56  
 ; CURRENT APPLICATION NUMBER: US/10/324,143  
 ; CURRENT FILING DATE: 2002-12-20  
 ; PRIOR APPLICATION NUMBER: 60/341,252  
 ; PRIOR FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 160  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 5  
 ; LENGTH: 816  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 ; OTHER INFORMATION: modified Streptococcus pneumoniae nucleotide  
 ; OTHER INFORMATION: sequence  
 US-10-324-143-5

Alignment Scores:  
 Pred. No.: 3.88e-65 Length: 816  
 Score: 822.50 Matches: 168  
 Percent Similarity: 72.83% Mismatches: 17  
 Conservative: 17

Best Local Similarity: 66.14% Mismatches: 28  
 Query Match: 20.47% Indels: 41  
 DB: 16 Gaps: 5

US-09-765-272A-66 (1-763) x US-10-324-143-5 (1-816)

```

QY 498 GlnIleThrTyrThrAspAspGluLeuGlnValAlaLysLeuAlaGlySerThrThr 517
Db 4 CAATATTACCTACATGATGAGATTGAGTTCAGGAGCAAGTTCGAGGCAAGTACACACA 63
QY 518 GluAspGlyTyrIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrVal 537
Db 64 GAAGACGGTTATATCTTGAT----- 84
QY 538 ThrProHisMetThrHisSerHisTrpLeuLysAspSerLeuSerGluAlaGluArg 557
Db 85 -----ACTAGTTGGATTAAAGATAGTTTCTCTGAAGCTGAGAGA 126
QY 558 AlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGln 577
Db 127 CGCGCAGCCCGGCTTATGCTAAAGAGAAAGGTTTGAACCCCTCTTCGACAGACCACCA 186
QY 578 AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLysAlaAla 597
Db 187 GATTGAGGAATACATGAGGCAAAAGAGCAGAGCTATCTACACCGCTGAAGCAGCT 246
QY 598 LysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAsn 617
Db 247 AAGAAGGTGCCACTTGATCGTATGCTTACATCTTCAATCTTCACTGATAGTCAAAAC 306
QY 618 GlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAsp 637
Db 307 GGTAGTTTAATCATACCTCATCATGACCATACCATACATCAATCAATTTGAGTGGTTGAC 366
QY 638 GluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLys 657
Db 367 GAAGCCCTTATGAGGACCTTAAGGGGTATAGTCTTGAGGATCTTTGGCGACTGTCAAG 426
QY 658 TyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSer 677
Db 427 TACTATGTCGAA-----COGCGG-----AACGCTAGT 453
QY 678 AspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSer 697
Db 454 GACCATGTTCTGTAATAAATAAG-----GCAGACCAAGATAGTAAACCTGAT 498
QY 698 GluGluLysProGlnThrGluLysProGluGluThrProArgGluGluLysProGln 717
Db 499 GAAGATAGGACATGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 558
QY 718 SerGluLysProGluSerPro-----LysProThrGluGluProGlu 731
Db 559 AATCACGCTGTTTAAATCCTTCAGCAGATAATCTTTATAACCAAGCACTGATACGGAA 618
QY 732 GluSerProGluGluSerGluGluProGlnValGluThrGlu 745
Db 619 GAGACAGAGGAAGAGTGAAGATACCAACACAGATGAGGCTGAA 660

```

RESULT 39

US-09-769-736-71  
 ; Sequence 71, Application US/09769736  
 ; Publication No. US20030138775A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Microbial Technics Limited  
 ; APPLICANT: Le Page, Richard WF  
 ; APPLICANT: Wells, Jeremy M  
 ; APPLICANT: Hanniffy, Sean B  
 ; TITLE OF INVENTION: Proteins  
 ; FILE REFERENCE: PWC/P21089wo  
 ; CURRENT APPLICATION NUMBER: US/09/769,736  
 ; CURRENT FILING DATE: 2003-02-14  
 ; PRIOR APPLICATION NUMBER: GB 9816335.5  
 ; PRIOR FILING DATE: 1998-07-27

; PRIOR APPLICATION NUMBER: US 60/125163  
 ; PRIOR FILING DATE: 1999-03-19  
 ; NUMBER OF SEQ ID NOS: 212  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 71  
 ; LENGTH: 1455  
 ; TYPE: DNA  
 ; ORGANISM: Streptococcus agalactiae  
 US-09-769-736-71

Alignment Scores: 2.04e-39 Length: 1455  
 Score: 542.50 Matches: 144  
 Percent Similarity: 34.29% Conservative: 59  
 Best Local Similarity: 24.32% Mismatches: 104  
 Query Match: 13.50% Indels: 285  
 DB: 13 Gaps: 13

US-09-765-272A-66 (1-763) x US-09-769-736-71 (1-1455)

```

QY 287 ThrAlaArgGlyValAlaValProHisGlyAsnHisTyrHisPheIleProTyrGluGln 306
Db 28 ACAGCAATCGGTGTTCTATTCCGACGGTAATCATTTCCACTTTATCTACTATAAGGAT 87
QY 307 MetSerGluLeuGluLysArgIleAlaArgIleLeuProLeuArgTyrArgSerAsnHis 326
Db 88 ATGCTCCATTAGAGTTAGAACCAACAAAGGATGGTG---GCAGAGCATAGAGGACATCAT 144
QY 327 TrpValProAspSerArgProGluGlnProSerProGlnSerThrProGluProSerPro 346
Db 144 ----- 144
QY 347 SerProGlnProAlaProAsnProGlnProAlaProSerAsnProIleAspGluLysLeu 366
Db 145 -----ATTGAT----- 150
QY 367 ValLysGluAlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSer 386
Db 150 ----- 150
QY 387 ArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGlyIleAspSerLysLeu 406
Db 151 -----GCATTAGGGAAAAAAGATCTTACA 174
QY 407 AlaLysGlnGluSerLeuSerHisLysLeuGlyAlaLysLysThrAspLeuProSerSer 426
Db 175 GAGAAACCAACGATATTTCTCATGAACCTAAATAAGGAA-----CCTCACACA 222
QY 427 AspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAlaArgIleHisGlnAspLeu 446
Db 223 GAGGAAGAACACCAT----- 237
QY 447 AspAsnLysGlyArgGlnValAspPheGluAlaLeuAspAsnLeuLeuGluArgLeuLys 466
Db 237 ----- 237
QY 467 AspVal***SerAspLysValLysLeuVal***AspIleLeuAlaPheLeuAlaProIle 486
Db 238 -----GCAGTAACA 246
QY 487 ArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTyrThrAspAspGluIle 506
Db 247 CGAAAGACCAACGTAAGGCAAAACCAATAGCCAGATTGCTACAGATTGCTCAAGAAATT 306
QY 507 GlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyrIlePheAspProArg 526
Db 307 GAAGAGGCAAAAAAGCTCGTAAATACACACATCTGATGGTTTACATTTTGTGCTAAA 366
QY 527 AspIleThrSerAspGluGlyAspAlaTyrValThrProHisMetThrHisSerHisTyr 546
Db 367 GATATTAAAAAGATACAGGTATGATTCAGGTTATGATTCATTCACATATGACATGAGCATGG 426
QY 547 IleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGlnAlaTyrAla----- 564

```

```
db 427 GTACCAAGAAAGATTATCAGAGTCGGAATTAAGACAGCTCAAGAAATTTCTTTCAGGA 486
2y 564 -----
db 487 AATCTGAAGCAATCAAGACAAACCAAAACAGGTAAACAGCTCAAGAAATCTATGAG 546
2y 564 -----
db 547 GCAATTGAACCAAAAGCAATTTGTAACCTGAAGATTATATTGGAATTGCACAGCG 606
2y 564 -----
db 607 ACAGACTATAAGATGGTACATTGTAATTCCTCATAAAGATCATTACCAATTATGTGAA 666
2y 565 -----
db 667 TTAATGTTGATGAAGAAAGATCTTTAGCTGATTCAGATAGACATATCTTTA 726
2y 569 -----
db 727 GAAGACTATTAGTACGCTAAATATTATCATGATGCACCCAGAAACGTCCTAAAGTT 786
2y 569 -----
db 787 GAAGATGGGTAAAGATGCTGAATTTATAAGAAAGGACTCTAATAAGCAGATAAA 846
2y 570 ---ThrProSerThrAspHisGlnAspSerGlyAsnThrGluAlaLysGlyAlaGlu 588
db 847 CCAAGTCTGCACCACTGATTAATAATCAACATCAATCTAGTCACAAAACCTTAAGT 906
2y 589 Ala-----IleTyrAsnArgValLysAlaLysLysValProLeuAspArgMetPro 606
db 907 GCTCAGAAAGTATCAACCAAGCAAAACCAAGAAATTTGTACCGCTTGATAAAATGCT 966
2y 607 TyrAsnLeuGlnTyrThrValGluValLysAsnGlySerLeuIlelleProHisTyrAsp 626
db 967 GCTCATGTCGATATGCAGTTGATTTGAAGATGATCAATTTGATTTCTCTCATCATGAT 1026
2y 627 HisTyrHisAsnIleLysPheGluTrpPheAspGlu---GlyLeuTyrGluAlaProLys 645
db 1027 CATATCATATATGTTCTTATGCTATGCTGTTTGACAGGGTGGTTATGGAAGCACCAGAA 1086
2y 646 GlyTyrThrLeuGluAspLeuAlaThrValLysTyrTyrValGluHisProAsnGlu 665
db 1087 GGTATACATTACCAACTCTTCTCAACAATTAATACTACATGGAACATCTTATGAA 1146
2y 666 ArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAsn 685
db 1147 TTACCA---AAAGAAAGGGTTGGGACACGACAGTGCAT-----AACAAAGGC 1194
2y 686 GlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLys 705
db 1195 TCAATAAAGCAATAAAGCAAAATATGCTCCAGATGAAGAACTGAA----- 1245
2y 706 ProGluGluGluThrProArgGluGluLysProGlnSerGluLysProGluSerProLys 725
db 1246 -----GATTGAGGAAA 1257
2y 726 ProThr-----GluGluProGlu 731
db 1258 GTAACCTCAACTATGTTTATGATGTTTAAAGGTTTCAGACCAAGAAAGAACAGAA 1317
2y 732 GluSerProGluGluSerGluGluProGlnValGlu 743
db 1318 AAACAAGAGATGAATCAGACTAGATGAATGAA 1353
2y 1318 AAACAAGAGATGAATCAGACTAGATGAATGAA 1353
db 1353 AAACAAGAGATGAATCAGACTAGATGAATGAA 1353
```

RESULT 40

US-09-765-599-34/c

Sequence 34, Application US/09452599

Patent No. US20020055101A1

GENERAL INFORMATION:

APPLICANT: Bergeron, Michel G.

APPLICANT: Ouellette, Marc

APPLICANT: Roy, Paul H.

TITLE OF INVENTION: Specific and Universal Probes and Amplification Primers

TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial

TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical

TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro

FILE REFERENCE: 12287.31

CURRENT APPLICATION NUMBER: US/09/452,599

PRIOR FILING DATE: 1999-12-01

PRIOR APPLICATION NUMBER: 08/526,840

PRIOR FILING DATE: 1995-09-11

PRIOR APPLICATION NUMBER: 08/304,732

PRIOR FILING DATE: 1994-09-12

NUMBER OF SEQ ID NOS: 177

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 34

LENGTH: 841

TYPE: DNA

ORGANISM: Streptococcus pneumoniae

US-09-452-599-34

Alignment Scores:

Pred. No.:	5,62e-37	Length:	841
Score:	512.00	Matches:	104
Percent Similarity:	93.10%	Conservative:	4
Best Local Similarity:	89.66%	Mismatches:	2
Query Match:	12.74%	Indels:	7
DB:	9	Gaps:	2

US-09-765-272A-66 (1-763) x US-09-452-599-34 (1-841)

QY 653 LeuAlaThrValLysTyrValGluHisProAsnGluAlaGluProHisSerAspAsnGly 672

Db 839 TTGGCGACTGTCAGTACTAGTCGAATCCAAACGACGTCG-CATTCAAGATATGTT 781

QY 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692

Db 780 TTTCGTAACGCTAGCGACCATGTTCCAAAGAAACAAATGGTCAAGCTGATACCAATCAA 721

QY 693 ThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluThrProArg 712

Db 720 ACGGAAAAACCCAAAGAGGAGAAACCTCAGACAGAAAAACCTGAGGAGAAACCCCTCGA 661

QY 713 GluGluLysProGlnSerGluLysProGluSerProLysProThrGluGlu----- 729

Db 560 GAAGAGAAACCGCAAGCGAGAAACCGAGAGTCTCCAAACCAACAGAGAACCCAGAGAA 601

QY 730 -----ProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLysVal 747

Db 600 GAATCACCAGAGAAATCACCAGAGGAATCAGAGAACCTCAGGTCGAGACTGAAAGTTA 541

QY 748 GluGluLysLeuArgGluAlaGluAspLeuGlyLysIleGlnAsp 763

Db 540 ---AGAAACTGAGAGAGGCTGAAGATTTACTTGGAAAAATCCAGAA 496

RESULT 41

US-10-121-120-34/c

Sequence 34, Application US/10121120

Publication No. US20030180733A1

GENERAL INFORMATION:

APPLICANT: Bergeron, Michel G.

APPLICANT: Ouellette, Marc

APPLICANT: Roy, Paul H.

TITLE OF INVENTION: Specific and Universal Probes and Amplification

TITLE OF INVENTION: Primers

TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial

TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical

TITLE OF INVENTION: Specimens for Routine Diagnosis in Micro

FILE REFERENCE: 12287.31

CURRENT APPLICATION NUMBER: US/10/121,120

PRIOR FILING DATE: 2002-04-11

PRIOR APPLICATION NUMBER: 09/452,599

PRIOR FILING DATE: 1999-12-01



; PRIOR APPLICATION NUMBER: 08/304,732  
; PRIOR FILING DATE: 1994-09-12  
; NUMBER OF SEQ ID NOS: 177  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 841  
; TYPE: DNA  
; ORGANISM: Streptococcus pneumoniae  
US-10-121-120-34

Alignment Scores:  
Pred. No.: 5,62e-37 Length: 841  
Score: 512.00 Matches: 104  
Percent Similarity: 93.10% Conservative: 4  
Best Local Similarity: 89.66% Mismatches: 2  
Query Match: 12.74% Indels: 7  
DB: 15 Gaps: 2

US-09-765-272A-66 (1-763) x US-10-121-120-34 (1-841)

```
QY 653 LeuAlaThrValLysTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGly 672
Db 839 TTGGCGACTGTCAAGTACTATGTGCAACATCCAAACGAAAGCTCG-CATTCAGATATGCT 781
QY 673 PheGlyAsnAlaSerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGln 692
Db 780 TTGGTAACGCTAGCGACCATGTTCCAAAGAAACAAATAATGGTCAAGCTGATACCAATCAA 721
QY 693 ThrGluLysProSerGluGluLysProGlnThrGluLysProGluGluLysThrProArg 712
Db 720 ACGGAAACCAACAGAGAGAAACCTCAGACAGAAACCTCAGAGAAACCCCTCGA 661
QY 713 GluGluLysProGlnSerGluLysProGluSerProLysProThrGluGlu 729
Db 660 GAACAGAAACCGCAAGCGAAACACAGAGTCTCCAAACCAACAGAGAGAAACCGAGAA 601
QY 730 -----ProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLysVal 747
Db 600 GAATCACAGAGAAATACCCAGAGAAATCAGAGAAACCTCAGGTCGAGACTGAAAGGTTA 541
QY 748 GluGluLysLeuArgGluAlaGluAspLeuLeuGlyLysIleGlnAsp 763
Db 540 ---AGAAACTGAGAGAGCTGAGATTACTTGGAAAAATCCAGAA 496
```

## RESULT 42

US-10-104-047-536  
; Sequence 536, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; PRIOR FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 536  
; LENGTH: 3466  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-104-047-536

Alignment Scores:  
Pred. No.: 5,34e-07 Length: 3466  
Score: 190.50 Matches: 167  
Percent Similarity: 33.55% Conservative: 97  
Best Local Similarity: 21.22% Mismatches: 302  
Query Match: 4.74% Indels: 221  
DB: 16 Gaps: 40

US-09-765-272A-66 (1-763) x US-10-104-047-536 (1-3466)

```
QY 74 LysValProTyrAspAla-----IleIleSerGluGluLeuLeuMetLysAspProAsnTyr 92
Db 1212 AAAGTGGCGGTAGATGTAATTAAGCAAGCTGTGGAGAATAATCAAGCCAATCTT 1271
QY 93 GlnLeuLysAspSerAspIleValAsnGluIle-----LysGlyGlyTyrValIle 109
Db 1272 GCATCCAAAGAACAGAAACAATCAATGCTTCAACCATACCAAGGTTGGCAT----- 1325
QY 110 LysValAsnGlyLysTyrTyrValTyrLeuLysAspAlaAlaHisAlaAspAsnIleArg 129
Db 1326 TCTGTTATTTCAGCCATCAGCTCTTCTTGGTTGATCAAGATGGAACACCAAAATATATC 1385
QY 130 ThrLysGluGluLysArgGlnLysGlnGluArgSerHisAsnHisAsnSerArgAla 149
Db 1386 ATCAACTACAGTCTTGAGCAGCCTAGCCAACTTCAAGTTGTTCTCCTCAAAATTTAAAAA 1445
QY 150 AspAsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrAspAspGlyTyrIle 169
Db 1446 GAAATCCAGTCCGCTACAAACAGTTGTAAAGTGAAAGTTTACCAGAGATCTTACTCTT 1505
QY 170 PheAsnAlaSerAsp-----IleIle 176
Db 1506 AAGTCTGAGAGAGCAAAAGCTTTGAAGGGGGGTGAATGATAGCAGCTTGTCTTCTGTGT 1565
QY 177 GluAsp-----ThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIlePro 195
Db 1566 GATGATGTTCCAGAGATATTAATGCATCTCCAGAAATTAAGCACTATGACCTA----- 1619
QY 196 LysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLysGln 215
Db 1620 AAGCAGCCTACTCAGCCTCTCCTCAGTCCCTGAGCAGAGAGCT-----GAGAGCCTGAG 1673
QY 216 GlySerArgProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSer 235
Db 1674 TCTCTGTTTTCAGTACTCTGGGAGTGGCAATTTGCTCTCTAGTCAGCAGCCTTTA--- 1730
QY 236 GluAsnHisAsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyLysIleSer 255
Db 1731 -----AAGAACCTCTGTG 1742
QY 256 SerLeuLeuArgGluLeuTyr-----AlaLysProLeuSerGluArgHisValGlu 272
Db 1743 TCTCTCTAAAGCATATTTATGCTTTGAATGCACAAACCAAGTCAGAGAGAG----- 1793
QY 273 SerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAla 292
Db 1794 -----CTCTCAAAATTTGCTGATTCAGTAAC 1820
QY 293 ValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuLys 312
Db 1821 CTACCA-----CTGGATGTAGTAAAAAG 1844
QY 313 ArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAspSerArg 332
Db 1845 -----TGGTTTGAAGAGATGCAA 1862
QY 333 ProGluGlnProSerProGlnSerThrProGluProSerProSerProGln----- 349
Db 1863 GCTGGACAGATTTCAGTGCAGTCTTCTGAACCATCTTCTCTGAACCAAGCAAGTAAT 1922
QY 350 ---ProAlaProAsn-----ProGlnProAlaProSerAsn---ProIleAspGlu 364
Db 1923 ATCCCTGCCAAGAACATGATCAGCTCAATCTGCAAAATGCAATGACCCAGGAGCAGC 1982
QY 365 -----LysLeuValLysGluAlaValArgLysValGlyAsp 376
Db 1983 ACAGTAAATCTACAAGTCTTTTGAAGATGACTAACTCCCACTTTTACCAGTGGGATCA 2042
QY 377 GlyTyrValPheGluGluAsnGlyValSerArgTyrIleProAla---LysAspLeuSer 395
Db 2043 -----ACCACCAATGGTTTCCAGAGAGTAGTACACCATCCCAATCACCTCTTAAC 2090
```



177	QY	177	GlUAsp---ThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIlePro	195
1606	Db	1606	GATGATTGTCAGAGAGATATTAAATGCACTTCCGAAATAAAGCACTATGACCTA	1659
196	QY	196	LysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLysGln	215
1660	Db	1660	AGACGCCTACTCAGCTCTCCACTCCCTGCGAGCAGAACT-----GAGAACGCTGAG	1713
216	QY	216	GlySerArgProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSer	235
1714	Db	1714	TCCTCTGTTTCATCAGCTACTCGAGATGGCAATTGTCTCTAGTCAGCACCTTTA	1770
236	QY	236	GluAsnHisAsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSer	255
1771	Db	1771	-----AAGAACCCTCTTG	1782
256	QY	256	SerLeuLeuArgGluLeuTyr-----AlaLysProLeuSerGluArgHisValGlu	272
1783	Db	1783	TCCTCTCTAAAGCATATTATGCTTTGAATGCACACCAAGTCGAGAAAGAG-	1833
273	QY	273	SerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAla	292
1834	Db	1834	-----CTCTCAAAATTCGTGATTCAGTAAAC	1860
293	QY	293	ValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetserGluLeuGluLys	312
1861	Db	1861	CTACCA-----CTGGATGTAGTAAAAAAG	1884
313	QY	313	ArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAspSerArg	332
1885	Db	1885	-----TGGTTGAAAAAGATGCAA	1902
333	QY	333	ProGluGlnProSerProGlnSerThrProGluProSerProSerProGln-----	349
1903	Db	1903	GCTGGACAGATTTCAGTCGAGCTTCTCGAACCATCTCTCTGACCGCAAGTAAAT	1962
350	QY	350	ProAlaProAsn-----ProGlnProAlaProSerAsn---ProfileAspGlu	364
1963	Db	1963	ATCCCTGCCAAGAACCAATGATCAGCTCTAATCTGCAATGCAAAATGAACCCAGGACAGC	2022
365	QY	365	-----LysLeuValLysGluAlaValArgLysValGlyAsp	376
2023	Db	2023	ACAGTAAATCTCAAAAGTCCTTTGAAGATGACCTAACTCCCGAGTTTACCAGTGGGATCA	2082
377	QY	377	GlyTyrValPheGluAsnGlyValSerArgTyrIleProAla---LysAspLeuSer	395
2083	Db	2083	-----ACCACCAATGTTCCAGAGTAGTACACCATCCCATCCCTCTTAAAC	2130
396	QY	396	AlaGluThrAlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLys	415
2131	Db	2131	CTTTCCTCATCCAGAAATACACAGGGTTACTTTGTACACAGCTGAGGGTGCAAGAAGAG	2190
416	QY	416	LeuGlyAlaLysLysThrAsp-----LeuProSerSerAspArgGluPheTyrAsnLys	433
2191	Db	2191	CCACAGTAGAACCTCTTGATCTTTCCTACTACCAAGCAACAGGAGAGATTATTAGAAAGG	2250
434	QY	434	AlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuAspAsnLysGlyArgGlnVal	453
2251	Db	2251	TCA-----ACTATCACTAGTGTTTACCAG-----ACAGTGTATTATTCTGTC	2292
454	QY	454	AppPheGluAlaLeuAspAsnLeuLeuGluArgLeuLysAspVal***SerAspLysVal	473
2293	Db	2293	CAGGAAGAACCCCTTG-----ACTTTGTCTTGGCGCAAAAAGGAGCCACAAAAGGAC	2343
474	QY	474	LysLeuVal**Asp-----IleLeuAlaPheLeuAlaProIleArgHisProGlu	490
2344	Db	2344	AGTTGTGTTACAGACTCAGAACCGATTGTAATGTAATCCCAAGTGCACACCCCAT	2403
491	QY	491	ArgLeuGlyLysProAsnAla-----GlnIleThrTyrThrAspAspGlu	505
2404	Db	2404	AATATCGCTATACCTACAGTCACTGCCAGTTACCCACCAATCTGGCCATTCGTGACCAG	2463
506	QY	506	IleGlnVal-----AlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyr	521

Db	2464	AACAGTGTTCATGCTTAAGAGCGCTAGCTGCCAATAAGCAACG	2508
Qy	522	IlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMet	541
Db	2509	-----ATTCTGATTCCCGAGTG	2526
Qy	542	ThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGln	561
Db	2527	GCATACACCTAC-----TCAACTACGGTCAGCCCTGCAGTCCAA	2565
Qy	562	AlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsn	581
Db	2566	-----GAACCAACCTTGAAAGTGATCCACGCCAAATGGAAT	2601
Qy	582	-----ThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLys	595
Db	2602	CAGATCAAGACACAGATACTAGCTCAGAAGGAGTATCAAAATGGAGGATCAGAATGAC	2661
Qy	596	AlaAlaLysLysValProLeuAspArgMet-----ProTyrAsnLeuGlnTyrThrVal	613
Db	2662	TCTGATTCTACACGCCCAAAAGAAATCGGAGACAGAAAATGGAATGTATGCTGT	2721
Qy	614	GluVal-----LysAsnGlySerLeuIleIleProHisTyrAspHis	627
Db	2722	GATTGTGTGACAAAGATATTTCCAAAAGAGTAGTTCATTATTGACACATAAATATGAACAC	2781
Qy	628	-----TyrHis	629
Db	2782	ACAGTAAAAGACCTCATGAGTGTGGAATCTGTAAGAAAGGCATTTAAACACAAACATCAT	2841
Qy	630	AsnIleLysPheGluTrpPheAspGluGlyLeuTyrGluAlaProLysGlyTyrThrLeu	649
Db	2842	TTGATTGAA-----CACATGCGATTACATTTCTGGAGAAAAGCCCTATCAATGT	2889
Qy	650	GluAspLeuLeuAlaThrValLysTyr-----TyrValGluHisProAsn-----	664
Db	2890	GACAAATGTGGAAAGCGCTTCTCACACTCTGGGCTCTTATTTCTCAACACATGAATCATCGC	2949
Qy	665	-----GluArgProHisSerAspAsnGlyPheGlyAsn	675
Db	2950	TACTCTCTACTGTAAGAGAGAAGCGGAAGACGTGACAGCACAGCAGGAGGAGGCGAGG	3009
Qy	676	-----AlaSerAspHisVal-----GlnArgAsnLysAsnGlyGlnAlaAspThr	690
Db	3010	CCTGAAATCCTCGAATGATGACAGTGGGTGCGCAGGGCGTCTCCCTCACAGGCGCACTCG	3069
Qy	691	AsnGlnThrGluLys---ProSerGluGluLysProGlnThrGluLysProGluGluGlu	709
Db	3070	GACGAGAGAGAGAGTTGACACAGGGAGAGGATGAACACAGTGAAGAAAGAGGAGGAG	3129
Qy	710	ThrProArgGluGluLysProGlnSerGluLysProGluSerProLysProThrGluGlu	729
Db	3130	GAGGATAAAGAGATGGAGAAATTCAGGAAGAAGAAGATGTCAAAAACCAACAGGGGAT	3189
Qy	730	ProGluGluSerProGluGluSerGluGluProGlnValGluThrGluLysValGluGlu	749
Db	3190	GAGGAA-----GAGGAGGAGGAGGAGGAAGAAGTGGAGAAGAAGAGGTAGAAGAG	3240
Qy	750	LysLeuArgGluAlaGluAsp	756
Db	3241	GCAGAGAATGAGGGAGAGAA	3261

RESULT 44

US2004-159-563-295  
; Sequence 295, Application US/10159563  
; Publication No. US20040009154A1  
; GENERAL INFORMATION:  
; APPLICANT: Khan, Javed  
; APPLICANT: Rignier, Markus  
; APPLICANT: Peterson, Carsten  
; APPLICANT: Meltzer, Paul  
; TITLE OF INVENTION: SELECTIONS OF GEN

TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS

FILE REFERENCE: 11613.56US11

CURRENT APPLICATION NUMBER: US/10/159,563

CURRENT FILING DATE: 2002-12-09

PRIOR APPLICATION NUMBER: US 10/133,937

PRIOR FILING DATE: 2002-04-25

NUMBER OF SEQ ID NOS: 444

SOFTWARE: PatentIn version 3.1

SEQ ID NO 295

LENGTH: 5316

TYPE: DNA

ORGANISM: Homo sapiens

US-10-159-563-295

Alignment Scores:

Pred. No.: 2,21e-06 Length: 5316  
Score: 186.50 Matches: 166  
Percent Similarity: 33.55% Conservative: 98  
Best Local Similarity: 21.09% Mismatches: 302  
Query Match: 4.64% Indels: 221  
Gaps: 40

US-09-765-272A-66 (1-763) x US-10-159-563-295 (1-5316)

74 LysValProTyrAspAla---lleleSerGluGluLeuLeuMetLysAspProAsnTyr 92  
1252 MAAGTGGCGGTAGATGCTAATGTAATGAAGCAAGTGTGGAGAATAATCAAGCCAACTCTT 1311  
93 GlnLeuLysAspSerAspIleValAsnGluLeu-----LysGlyGlyTyrValle 109  
1312 GCATCCCAAGAGAACAAACAATCATGCTTCCACCATACAAAGGTGGCCAT----- 1365  
110 LysValAsnGlyLysTyrTyrValTyrLeuLysAspAlaAlaHisAlaAspAsnIleArg 129  
1366 TCTGTTATTTCAGCCATAGTCTTCTCTTGGTTGATCAAGATCGAACCAAAATATATC 1425  
130 ThrLysGluGluLysArgGlnLysGlnGluArgSerHisAsnHisAsnSerArgAla 149  
1426 ATCACTACAGTCTTGAGCGCTAGCCACTTCAAGTTGTTCTCAAATTTAAAAA 1485  
150 AspAsnAlaValAlaAlaAlaArgAlaGlnGlyArgTyrThrAspAspGlyTyrIle 169  
1486 GAAATCCAGTCGCTACAAACAGTTGTAAAGAGTAAAGTACAGAGATCTTACTGTT 1545  
170 PheAsnAlaSerAsp-----llelle 176  
1546 AAGTCTGAGAGGACAAAAGCTTTGAAGGGGGGTGAATGATAGCACTTCTCTCTGTGT 1605  
177 GluAsp---ThrGlyAspAlaTyrIleValProHisGlyAspHisTyrHisTyrIlePro 195  
1606 GATGATGTCCAGGAGATATTAAATGCATCTCCAGATTAAAGCACTATGACCTA----- 1659  
196 LysAsnGluLeuSerAlaSerGluLeuAlaAlaGluAlaTyrTrpAsnGlyLysGln 215  
1660 AAGCAGCTTACTCAGCTCTCCACTCTCCCTCGCAGCAGAGCT-----GAGAAGCCTGAG 1713  
216 GlySerArgProSerSerSerSerTyrAsnAlaAsnProAlaGlnProArgLeuSer 235  
1714 TCTCTGTTTCATCAGCTACGTGGAGATGCAATTTCTCTCTAGTCAGCCACCTTTA--- 1770  
236 GluAsnHisAsnLeuThrValThrProThrTyrHisGlnAsnGlnGlyGluAsnIleSer 255  
1771 -----AGAACCCTCTTG 1782  
256 SerLeuLeuArgGluLeuTyr-----AlaLysProLeuSerGluArgHisValGlu 272  
1783 TCTCTCTAAAGCATATTATGCTTTGAATGACACCAACCAAGTCAGAGAG----- 1833  
273 SerAspGlyLeuIlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAla 292  
1834 -----CTCTCAAAATTTGCTGATTTCAGTAAAC 1860  
293 ValProHisGlyAsnHisTyrHisPheIleProTyrGluGlnMetSerGluLeuGluLys 312

1861 CTACCA-----CTGATGTAGTAAAAAG 1884  
313 ArgIleAlaArgIleIleProLeuArgTyrArgSerAsnHisTyrValProAspSerArg 332  
1885 -----TGGTTTGAAGATGCAA 1902  
333 ProGluGlnProSerProGlnSerThrProGluProSerProSerProGln----- 349  
1903 GCTGGACAGATTTCAGTGCAGTCTTCTGAACCATCTTCTCTGAACCAAGGCAAAAGTAAAT 1962  
350 ---ProAlaProAsn-----ProGlnProAlaProSerAsn---ProIleAspGlu 364  
1963 ATCCCTGCCAAGAACAAATGATCAGCCTCAATCTGCAAAATGCAAAATGAACCCGAGGACAGC 2022  
365 -----LysLeuValLysGluAlaValArgLysValGlyAsp 376  
2023 ACAGTAATCTACAAAGTCCTTTGAGATGACTAATCCCAAGTTTACCACTGGATCA 2082  
377 GlyTyrValPheGluGluAsnGlyValSerArgTyrIleProAla---LysAspLeuSer 395  
2083 -----ACCACCAATGGTTCAGAGAGTAGTACACCATCCCATCATCCTCTAAAC 2130  
396 AlaGluThrAlaAlaGlyIleAspSerLysLeuAlaLysGlnGluSerLeuSerHisLys 415  
2131 CTTTCTCATCCAGAAATACACAGGGTTTACTTGTACACAGCTGAGGGTGCACAAAGAGAG 2190  
416 LeuGlyAlaLysLysThrAsp-----LeuProSerSerAspArgGluPheTyrAsnLys 433  
2191 CCACAAAGTAGAACCTCTTGATCTTTCTACTACCAAGCAACAGGAGAAATATTAGAAAGG 2250  
434 AlaTyrAspLeuLeuAlaArgIleHisGlnAspLeuLeuAspAsnLysGlyArgGlnVal 453  
2251 TCA-----ACTATCCTACTAGTGTTCACAG-----AACAGTCTTATTCTGTC 2292  
454 AspPheGluAlaLeuAspAsnLeuLeuArgLeuLysAspVal\*\*\*SerAspLysVal 473  
2293 CAGGAAGAACCCCTTG---AAGTGTCTTTCGCAAAAGAGGAGCCACAAAGAGAC----- 2343  
474 LysLeuVal\*\*\*Asp-----lleleAlaPheLeuAlaProIleArgHisProGlu 490  
2344 AGTTGTGTACAGACTCAGAACCACTGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 2403  
491 ArgLeuGlyLysProAsnAla-----GlnIleThrTyrThrAspAspGlu 505  
2404 AATATCGTATACCTACAGTCACTGCCAGTTCACCAATCGTGGCCATTGCTGACCAG 2463  
506 IleGlnVal-----AlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTyr 521  
2464 AACAGTGTTCATGCTTAAGAGCGCTAGTGCATAAAGCAACG----- 2508  
522 IlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMet 541  
2509 -----ATTCTGATCCCCAGGTG 2526  
542 ThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaGln 561  
2527 GCATACACCTAC-----TCAACTACGCTCAGCCTCGAGTCCAA 2565  
562 AlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGlnAspSerGlyAsn 581  
2566 -----GAACCACTTGAAGTATCCAGCAAAATGGAAT 2601  
582 -----ThrGluAlaLysGlyAlaGluAlaIleTyrAsnArgValLys 595  
2602 CAGGATGAAGACAAAGATACCTAGTCTCAGAGAGTATCAATGTAGAGGATCAGATGAC 2661  
596 AlaAlaLysLysValProLeuAspArgMet-----ProTyrAsnLeuGlnTyrThrVal 613  
2662 TCTGATTCTACCGCCCAAAAGAAATGCGGAGACAGAAAAATGGAATGATGCTGTTGT 2721  
614 GluVal-----LysAsnGlySerLeuIleIleProHisTyrAspHis 627











```
QY 524 AspProArgAspIleThr-----SerAspGluGlyAspAlaTyrVal 537
Db 2488 TCCTCTCTGTTGCTACTAAAGGACCTCTTCTCAATGAAGAGAAAT----- 2538
QY 538 ThrProHisMetThrHisSerHisTyrIleLysLysAspSerLeuSerGluAlaGluArg 557
Db 2539 -----CGTAAATGTTAAACGACGACCAAGAT 2568
QY 558 AlaAlaAlaGlnAlaTyrAlaLysGluLysGlyLeuThrProProSerThrAspHisGln 577
Db 2569 GAGTCTCGCAGATGTAGTCAGCAGCAAGGAAACAACTTCTAGGTTGTCAGATTGAA 2628
QY 578 AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaLysLeuThrAsnArgValLysAlaAla 597
Db 2629 ---AGCCCAACATA-----GGAGGTGAAGACTTGAACAAGAACGACGAGGGTCC 2676
QY 598 LysLysValPro-LeuAspArgMetProTyrAsnLeuGlnTyrThrValGluValLysAs 617
Db 2677 AAGGAA---CCAGAAAGCATCTGCCAATATACAGGATGCTAAGATAAATAAATACCTG 2733
QY 617 nGlySerLeuIleProHisTyrAspHisTyrHisAsnIleLysPheGluTyrPheAs 637
Db 2734 AAGGAAAGGAGACTCCAAAGCAACACAGGTACT-CAGAAAAAGTTAAATGACAGA 2792
QY 637 p-----GluGlyLeuTyrGluAlaProLys-----GlyTyrThrLeuGluAs 651
Db 2793 TCCTTAAAGAGGCAATGTTACAGTCTTCTTAAGCCAGTCAATGATAAATGAAGGA 2852
QY 651 pLeuLeuAla-----ThrValLysTyrTyrValGluHisPro-----AsnGln 665
Db 2853 TTTCAGGCTAAATTTACTAATAGTAAACACACTCGTGCCTCGAATAAATTTCTAATGA 2912
QY 665 uArgProHisSerAspAsnGlyPheGlyAsnAlaSerAspHisValGlnArgAsnLysAs 685
Db 2913 AAAA-----AATAATCTAGTCABACGATGATCTTCTAGCTCCTCAGCTTCTTC 2963
QY 685 nGlyGlnAlaAspThrAsnGlnThrGluLysProSerGluGluLysProGlnThrGluLys 705
Db 2964 CACAGAAAGCAGTCTAGTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3023
QY 705 sProGluGlu-----GluThrProArgGluGluLysProGlnSerGluLysPr 721
Db 3024 GAAGGTAGNAGAGTGTGTTAAATACCAAGAGAACCCAGTTAGATCTTCTCATCATGAT 3083
QY 721 oGluSerProLysProThr-----GluGluProGluGluSerPro-- 734
Db 3084 TGAGGCCCTCTCTCCATCTCTTTAAACAAAATCAATGCCACCTCATATAAATCCAGT 3143
QY 735 -----GluGluSerGluGluProGlnValGluThrGlu----- 745
Db 3144 TACACAGCTAATGGACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3203
QY 746 -----LysValGluGluLysLeuArgGluAlaGluAs 756
Db 3204 GTCTAGTATCTACCTGACCTTCCAGAAAGTTGCTCTCTCTCTCTCTCTCTCTCT 3263
QY 756 pLeuLeuGlyLys 760
Db 3264 TCTTGTTTCAAGA 3276
```

RESULT 48  
US-09-842-552-98  
; Sequence 98, Application US/09842552  
; Patent No. US20020055628A1  
; GENERAL INFORMATION:  
; APPLICANT: The Regents of The University of California  
; TITLE OF INVENTION: MULTILOCUS REPETITIVE DNA SEQUENCES FOR GENOTYPING BACILLUS ANTHR  
; FILE REFERENCE: S-89,687  
; CURRENT APPLICATION NUMBER: US/09/842,552  
; CURRENT FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/199,911  
; PRIOR FILING DATE: 2000-04-26

; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 98  
; LENGTH: 2522  
; TYPE: DNA  
; ORGANISM: Bacillus anthracis  
US-09-842-552-98  
Alignment Scores:  
Pred. No.: 4, 7e-06 Length: 2522  
Score: 178.00 Matches: 119  
Percent Similarity: 34.06% Conservative: 52  
Best Local Similarity: 23.71% Mismatches: 151  
Query Match: 4.43% Indels: 180  
DB: 9 Gaps: 27

US-09-765-272a-66 (1-763) x US-09-842-552-98 (1-2522)  
QY 382 GluAsnGlyValSerArgTyrIleProAlaLysAspLeuSerAlaGluThrAlaAlaGly 401  
Db 12 GAAATGCTCAATTAAATTTATGAAAAAAGATAGTGAATCTGGTCAACTTCTACCAAGT 71  
QY 402 Ile-----AspSerLysLeuAla----- 407  
Db 72 GCAAAATTTGATGTTATCGATAAAGATGGAAAGTTGTGAAACAATTTGTTACAGATGAT 131  
QY 408 LysGlnGluSerLeuSerHisLysLeuGlyAla-----LysLysThrAsp 422  
Db 132 AAAGTGAAGCTTTATCGAAACAACCTCCAGTTGGAAGCTATACATTAAAGAAAGTAGAA 191  
QY 423 LeuProSerSerAspArgGluPheTyrAsnLysAlaTyrAspLeu--LeuAlaArgIle 441  
Db 192 GCACCG-----AAAGGATATGAATTATCACTAGTTCAGTT 227  
QY 442 HisGlnAspLeuLeuAspAsnLysGlyArgGlnValAspPhe----- 455  
Db 228 TCTGTTGATGTAGAGGCTAAATAAAGTAGTACTGTAGTGTGTGTAATAAAGATCCCC 287  
QY 456 GluAlaLeuAspAsnLeuGluArgLeuLysAspVal\*\*SerAspLysValLysLeu 475  
Db 288 GAAAGAGTAACAGGTCAATTTGAAGTGTGTAAGTAGATGCAAAATGATAAAGCAAAATTG 347  
QY 476 Val\*\*AspIleLeuAlaPheLeuAlaProIleArgHisProGluArgLeuGlyLysPro 495  
Db 348 TTATCA----- 353  
QY 496 AsnAlaGlnIleThrTyrThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyr 515  
Db 354 GGTGCAAAATTCGAAGTGTATAAAGATGCAAAAAGGTAGCAGAACTGAAA----- 404  
QY 516 ThrThrGluAspGlyTyrIlePheAspProArg-----AspIleThrSerAsp 531  
Db 405 ACAGTGAGAGTGAAGAAAGTGTATGTCACGAAATTTACCGTAGGTGAATACACAGTAAA 464  
QY 532 GluGlyAspAlaTyrValThrProHisMetThrHisSerHisTyr----- 546  
Db 465 GAAACGAAAGCACCGCGGCTACAGCTTTCAGATAAAGATGGAAGTAACAATTCAA 524  
QY 546 ----- 546  
Db 525 AACGAGAAAGAGTAGTAAAGTAGAGCGCAGAAAAACGAAAAAATCTTAGTTCTCTACAA 584  
QY 547 ---IleLysLysAspSerLeuSerGluAlaGluArgAlaAlaAlaGlnAlaTyr----- 563  
Db 585 ATTATTAATGGATGATAAAGATCAACGAAACGCTTAGCAGCGCGCGCAATTTACATTG 644  
QY 564 -----AlaLysGluLysLeuThrProProSerThrAspHisGln 577  
Db 645 AAAGATGTAAAGCAATGTTGTAAGAAAGAGAAATTACA-----ACAGAT----- 689  
QY 578 AspSerGlyAsnThrGluAlaLysGly-----AlaGluAlaIleTyrAsnArgVal 594  
Db 690 AAGTCTCGAACTGTTAAAGTAGACGAGCTTGTGCGGGGTGATATACCTTAGAAGAAACA 749

595 LysAlaAlaLysLysValProLeuAspArgMetProTyrAsnLeuGlnTyrThrValGlu 614  
750 AAAGCCGCCAGAGGTTATAAGCA-----TTAGAAGTACATCGAA 791  
615 ValLys----- 616  
792 GTAACGCTAGTACAAACGAAGTAGTAAACACAGACGCTGTGTAATGAAAGAGTGAAGAA 851  
617 -----AsnGlySerLeuIleProHisTyrAspHisTyrHisAsnIleLys----- 632  
852 GAATACAGGCAATTAGAAATTACAAAGTAGATGCTAATGATATAATAATAATAATA 911  
633 -----PheGlu-----Tyr----- 635  
912 GCAGGCGCAGTGTGAAATTTGGAAGACGGAACAAATCGATACATTAACATCAGAT 971  
636 -----PheAspGluGlyLeuTyr----- 641  
972 GAAATGGTAAAGCAATTTGGAAGACGGAACAAATCGATACATTAATTTAAAGGAAGTT 1031  
642 GluAlaProLysGlyTyrThrLeuGluAsp-----LeuLeuAlaThrVal 656  
1032 CAAGCCCGCAGAGGTTATGAGTTATCTGATAAGGAATCGAATTTCAGATTCTTAATCAA 1091  
657 LysTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAla 676  
1092 AAATTTGAAGTTGTAAGAACTTCAATTTACAAATTAATAAAGAAACGAAGGTCGAG 1151  
677 SerAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysPro 696  
1152 AATCCAGGCGCAAGAAACAGAAAGACCCAGGTGAA-----GAAACAGAAAGCCG 1199  
697 SerGlu-----GluLysPro-----GlnThrGluLysProGluGluThrProArg 712  
1200 GGTGAAGAAACAGAAAGCCAGGCGAAGAAACAGAAAGCCAGGTGAAGAAACA----- 1253  
713 GluGluLysPro-----GlnSerGluLysPro-----GluSerProLysProThrGlu 728  
1254 ---GAAAGCCAGGTGAGAAACAGAAAGCCAGGCGAAGAAACAGAAAGCCAGGTGA 1310  
729 GluProGluLysSerProGluGluSerGluGluProGlnValGluThrGluLysValGlu 748  
1311 GAAACAGAAAGCCAGGTGAAGAAACAGAAAGCCAGGCGAAGAAACAGAAAGCCAGGT 1370  
749 GluLys 750  
1371 GAAGAA 1376

RESULT 49  
US-09-816-669A-13  
; Sequence 13, Application US/09816669A  
; Patent No. US20020137019A1  
; GENERAL INFORMATION:  
; APPLICANT: GARABEDIAN, Michael  
; APPLICANT: TANEJA, Samir  
; APPLICANT: HITTELMAN, Adam  
; APPLICANT: MARKUS, Steven  
; TITLE OF INVENTION: METHOD FOR SCREENING TRANSCRIPTIONAL COREGULATORY PROTEINS OF  
; TITLE OF INVENTION: TRANSCRIPTION FACTORS, AND ANDROGEN RECEPTOR TRANSCRIPTIONAL CO  
; TITLE OF INVENTION: PROTEINS AS TARGETS FOR ANDROGEN RECEPTOR-DEPENDENT DISEASES  
; FILE REFERENCE: GARABEDIAN-1, 1A  
; CURRENT APPLICATION NUMBER: US/09/816, 669A  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: 60/225,618  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/191,768  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 8588  
; TYPE: DNA

ORGANISM: Human  
US-09-816-669A-13  
Alignment Scores:  
Pred. No.: 2,78e-05 Length: 8588  
Score: 177.50 Matches: 138  
Percent Similarity: 30.01% Conservative: 70  
Best Local Similarity: 19.91% Mismatches: 201  
Query Match: 4.42% Indels: 285  
DB: 9 Gaps: 30  
US-09-765-272A-66 (1-763) x US-09-816-669A-13 (1-8588)  
QY 216 GlySerArgProSer-----SerSerSerSerTyrAsnAlaAsn 228  
DB 3163 GGTCTGTGACCTTCCATGCTCCTCCACGATGCTTTCAGATGCCAGAGGACGCCACC 3222  
QY 229 ProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThrTyrHisGln 248  
DB 3223 AAGGCC-----CTGCTGGAGAACTTTGGCTTTGAGTTGGTTCATCCAGTATATAG 3273  
QY 249 Asn-----GlnGlyGluAsnIleSerSer 256  
DB 3274 AACAGCAGAGGTGCAGAAAAAGAAATGGGAAGACTGACCCAGGAGAGAACCTTGGAAAAAG 3333  
QY 257 LeuLeuArgGluLeuTyrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeu 276  
DB 3334 CTC-----GAGTGTGACTCCTCGCGGCAAG 3357  
QY 277 IlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGly 296  
DB 3358 TTGTTTCCAACTCTTGATTTTAAAG-----AGTCATCAA 3393  
QY 287 AsnHisTyrHis-----PheIleProTyrGluGlnMetSerGlnLeuGluLysArgIle 314  
DB 3394 GAGCAGCTTTCATCAGATTTACTTTCTTCAACACAGCTCGAGAGGTTTGCCTCAACAGTAC 3453  
QY 315 Ala-----ArgIleIleProLeuArgTyrArgSerAsnHisTrpValProAsp 330  
DB 3454 AGAGACCACTAGATAAACTGTACCCTGAGGCGCCAGACC-----CCAGAG 3501  
QY 331 SerArgProGluGlnProSerProGlnSerThrProGluProSerProSerProGlnPro 350  
DB 3502 CCACCACCACTCCCTCCCTCCACCTCCACCCCTCCACCTCCGCGCAGCGCGCTCAGCG 3561  
QY 351 AlaProAsnProGln-----ProAlaProSerAsnProIleAspGluLysLeuValLysGlu 369  
DB 3562 GCGTCCACACCACTCCCTCCCTCCACCTCCACCTCCACCTCCACCTCCACCTCCACCT 3621  
QY 370 AlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSerArgTyrIle 389  
DB 3622 GCCCAGCATCAGTCCC-GCTCACCAGCTCTCCATGCCGATGGAGCTGCCCATCTCTC 3680  
QY 390 ProAlaLysAspLeu-----SerAlaGluThrAlaAla 400  
DB 3681 GCCGCTGATGTCAGACGATGCCGTCGACACCTTCCGCGCTCAGCTACCTCCCGCAGCT 3740  
QY 401 GlyIle----- 402  
DB 3741 GGGACCTGTGGAGCCTCTGCTGCGGACCTGGCCAACTCTTACCAGCATCAGCTCAATCC 3800  
QY 402 ----- 402  
DB 3801 AACCTGTCCAGCAGCAGAAACAGAGGCTTCGCCACAGGATCAGATGATCAGCTCCG 3860  
QY 402 ----- 402  
DB 3861 AGCTTTCGCGCAATATTTTGACATTAACAACTCCCGCTGAGAGCAATAAAGAGAT 3920  
QY 403 -----AspSerLysLeuAlaLysGlnGluSerLeu 412  
DB 3921 GGCAGACAAGTCCGGGTGCCCCCAGAAAGTATCAGACACTGGTTCAGGAACACTCTCTT 3980

QY	413	Ser	-----HisLys	415
DB	3981	CAAAGAGAGCGACGTAACAAGGACTCCCTTAACTTCAGTAATCTCTATCACCAG	4040	
QY	416	LeuGlyAlaIylsIylsThrAspLeuProSer	-----	425
DB	4041	CTGTGGAGGAGCTCAGATGTGACTCCCGGCCCTTCGCCGGAACCTCCAAAGCAGAGTA	4100	
QY	426	-----SerAspArgGluPheTyrAsnLysAlaIylsAspLeuLeuAla-ArgIleHisG	443	
DB	4101	CTGGGAAGCAAGAGGTCTTCAAGAACAGGTTTACGGACTACCACTGAGGCTCTTACA	4160	
QY	443	nAspLeuLeuAspAsnIylsGly-----ArgGlnValAspPheGluAlaLeuAspAsnIle	461	
DB	4161	GGACTTCTTCGATGCCAATCTTACCACAAAGGATGATGAATTTTGACCACTCTCTAATTT	4220	
QY	461	uLeuGluArgLeuLysAspVal***SerAspLysValLysLeuVal**AspIleLeuAl	481	
DB	4221	ACTGAACCTTCCAAACCGAGTG-----AYATGGT	4250	
QY	481	aPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTy	501	
DB	4251	GTGGTTTCAGAATGCCGACAGAGAAGCCAGGAAGATTATGAGAAT-----	4296	
QY	501	rThrAspAspGluIleGlnValAlaIylsLeuAlaGlyIylsTyrThrThrGluAspGlyTy	521	
DB	4297	-----CAGGAGAGAGGCCAAA-----GATGGA--	4317	
QY	521	rIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMe	541	
DB	4318	-----GACGGCGTGGCTTCAATGAT-----	4341	
QY	541	tThrHisSerHisTrpIleLysIylsAspSerLeuSerGluAlaGluArgAlaAlaG	561	
DB	4342	-----AGATACATTCGACCAACCACTGAATCACCAGTCGCAAAAATGTAGCCT	4391	
QY	561	nAlaTyAlaLys-----GluLysGlyLeuThrProProSerThr	574	
DB	4392	GGTGTTCAGCGCATCTTTGATCTCAAGCACCAGAGAAGCTGTGTACCAAGATGA	4451	
QY	574	rAspHisGln-----AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTy	591	
DB	4452	GGATGAGGAGGGCGAGCACACCAAAATGAGGATTCATGGATGCCATGGAAATCCT	4511	
QY	591	rAsnArgValLysAlaLalysLysValProLeuAspArgMetProTyrAsn-----	608	
DB	4512	GAGCGCTACAGCTCATCTCGTAGTACCCCGATGCCCTCACAGGCTTACAGCCCCCAGC	4571	
QY	609	-----LeuGlnTyrThrValGluValLysAs	617	
DB	4572	ACCATCAGCCAAATAATACAGCTTCTCCGCTTCTTTCAGCTTACAGCGGAGGTGAGGA	4631	
QY	617	nGlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleIylsPheGluTrpPheAs	637	
DB	4632	ACTGGCCACCTTC-----AATTCAAAACAGACGACGCGCA	4667	
QY	637	pGluGlyLeuTyTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValTy	657	
DB	4668	T-----GAGAAACCAAG-----	4680	
QY	657	sTyTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSe	677	
DB	4681	-----CTGGCGGAAGCTCCAGTGCACAGCCAAACCA-----	4713	
QY	677	rAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSe	697	
DB	4714	-----ACCCAGAAAAAGCAAGACAA-----CCAAA	4739	
QY	697	rGluGluLysProGlnThrGluLysProGluGluGlu-----ThrProArgGluGluTy	715	
DB	4740	GCCAGAGCTGCAGCAGCAGACAGACCGCCGAGAGAACCAACACTCCCCAGCAGAGCT	4799	
QY	715	sProGln-----SerGluLysProGluSerProLysPro-----	726	

```

Db      4800 CCCCAGCTGGTGTCCCTGCCTTCGTTGCCACAGCCTCTCCACAGCGCCCTCCACA 4859
          |||||
Qy      727 -ThrGlulgluProgludluSerProgluGlulSerGlu 738
          |||:::
Db      4860 GTGCCCCATTACCAGTAGTGAGCCCAAGCTCCTCCAG 4896
          |||:::

RESULT 50
US-10-342-887-1335
; Sequence 1335, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIORITY FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1335
; LENGTH: 8588
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1335

Alignment Scores:
Pred. No.:      2,78e-05      Length:      8588
Score:          177.50       Matches:     138
Percent Similarity: 30.01%    Conservative: 70
Best Local Similarity: 19.91% Mismatches:    201
Query Match:      4.42%      Indels:      285
DB:              13         Gaps:        30

US-09-765-272A-66 (1-763) x US-10-342-887-1335 (1-8588)
Qy      216 GlySerArgProSer-----SerSerSerSerTyrrAsnAlaAsn 228
          |||||
Db      3163 GGTTCTGAGCGTTCCATGCTCCCTCCACGCATTGCTTCAGATGCCAGAGGAACGCCACC 3222
          |||||
Qy      229 ProAlaGlnProArgLeuSerGluAsnHisAsnLeuThrValThrProThrTyrrHisGln 248
          |||||
Db      3223 AAGGCC-----CTGCTGGAGAAGACTTGGCTTTGAGTTGGTCATCATCAGTATAATGAG 3273
          |||||
Qy      249 Asn-----GlnGlyGluAsnIleSerSer 256
          |||||
Db      3274 AACAAAGCAGAGGTGCAGAAAAAGAAATGGGAAGACTGACCAAGGAGAGAACCTGGAAAAG 3333
          |||||
Qy      257 LeuLeuArgGluLeuTyrrAlaLysProLeuSerGluArgHisValGluSerAspGlyLeu 276
          |||||
Db      3334 CTC-----GAGTGTGACTCTCGCGCAAG 3357
          |||||
Qy      277 IlePheAspProAlaGlnIleThrSerArgThrAlaArgGlyValAlaValProHisGly 296
          |||||
Db      3358 TTGTTTCCAAATCTGATTTTAAG-----AGTCATCAA 3393
          |||||
Qy      297 AsnHisTyrrHis-----PheIleProTyrrGluGlnMetSerGluLeuGluLysArgIle 314
          |||||
Db      3394 GAGCACGTTTCATCAGAATTAATCTTCCTTTCAACAGCTCGAGAGGTTTGCCAACAGTAC 3453
          |||||
Qy      315 Ala-----ArgIleIleProLeuArgTyrrArgSerAsnHisTrpValProAsp 330
          |||||

```

3454 AGAGACCCTACGATAAATCTTACCCACTGAGGCCCCAGACC-----CCAGAG 3501  
331 SerArgProGluGlnProSerProGlnSerThrProGluProSerProSerProGlnPro 350  
3502 CCACCACCACTCCCTCCCTCCACCCCTCCACCCCTCCCTCCGCGAGCGCGCTCAGCG 3561  
351 AlaProAsnProGln---ProAlaProSerAsnProIleAspGluLeuValLysGlu 369  
3562 GCGTCCACACCAAGCATCCCGCGCATCGCCGCCACCCATCACCTACCAATTGCACCG 3621  
370 AlaValArgLysValGlyAspGlyTyrValPheGluGluAsnGlyValSerArgTyrIle 389  
3622 GCCAGCCATCAGTGCC-GCTACCCAGCTCTCCATGCCGATGAGCTGCCCATCTTCTC 3680  
390 ProAlaLysAspLeu-----SerAlaGluThrAlaAla 400  
3681 GCCGCTGATGATCGAGCATGCCGCTGCAGACCTTGGCGGCTCAGCTACCCCGCAGCT 3740  
401 GlyIle----- 402  
3741 GGGACCTGTGAGCCTCTGCCTCGGACCTGGCCCACTCTTACCAGCATCAGTCAATCC 3800  
402 ----- 402  
3801 AACCTGTCTCCAGCAGCAGAAACAAGAGCCTTGCACACGATCACAGATGATCAGCTCCG 3860  
402 ----- 402  
3861 AGTCTTGGCGCAATATTTTGACATTAACTCCCTCCCTCCAGTGNAGAGCAATAAAGAGAT 3920  
403 -----AspSerLysLeuAlaLysGlnGluSerLeu 412  
3921 GGCAGACAAGTCCGGTGTGCCCAAGAGTGTATCAAGCACTGGTTTCAGGAACACTCTCT 3980  
413 Ser-----HisLys 415  
3981 CAAAGAGAGGCGAGCGGTAAACAAGGACTCCCTTACAACTTCAGTAATCCTCTATCACCAG 4040  
416 LeuGlyAlaLysLysThrAspLeuProSer----- 425  
4041 CCTGGAGGAGCTCAAGATTGACTCCCGGCCCTTCGGCGGNAACCTCCAAAGCAGAGTA 4100  
426 -----SerAspArgGluPheTyrAsnLysAlaTyrAspLeuLeuAla-ArgIleHisG1 443  
4101 CTGGGGAAGCAAGAGGTCTTCAAGAAAGGTTTACGGACTACCACTGAGGCTTCTTACA 4160  
443 nAspLeuLeuAspAsnLysGly-----ArgGlnValAspPheGluAlaLeuAspAsnLe 461  
4161 GGACTTCTTCGATGCCAATGCTTACCCAAAGGATGATGAATTTGAGCAACTCTCTAAATT 4220  
461 uLeuGluArgLeuLysAspVal\*\*\*SerAspLysValLysLeuVal\*\*\*AspIleLeuAl 481  
4221 ACTGAACCTTCCACCGAGTG-----ATAGTGGT 4250  
481 aPheLeuAlaProIleArgHisProGluArgLeuGlyLysProAsnAlaGlnIleThrTy 501  
4251 GTGGTTTCAGATGCCCGACAGAGCGCCAGGAGATTAATGAGAT----- 4296  
501 rThrAspAspGluIleGlnValAlaLysLeuAlaGlyLysTyrThrThrGluAspGlyTy 521  
4297 -----CAGGGAGAGGGCAAA-----GATGGA-- 4317  
521 rIlePheAspProArgAspIleThrSerAspGluGlyAspAlaTyrValThrProHisMe 541  
4318 -----GAGCGCGGTGAGCTTACAAATGAT----- 4341  
541 tThrHisSerHisTrpIleLysLysAspSerLeuSerGluAlaGluArgAlaAlaG1 561  
4342 -----AGATACATTCGAACAGCAACTTGAATCAGTCCAAATAATGTAGCCT 4391  
561 nAlaTyzAlaLys-----GluLysGlyLeuThrProProSerTh 574  
4392 GGTGTTTCAGCGCATCTTTGATCTCATCAAGCACCAGGAAGAGTGTGTGTACAGGATGA 4451

QY 574 rAspHisGln-----AspSerGlyAsnThrGluAlaLysGlyAlaGluAlaIleTy 591  
Db 4452 GGATGAGGAGGGCGAGGACGACGCAAAATGAGGATTCATGATGCGATGAAATCCT 4511  
QY 591 rAsnArgValLysAlaAlaLysLysValProLeuAspArgMetProTyrAsn----- 608  
Db 4512 GAGCGCTACAGCTCATCTCTGAGTACCCCGATGCCCTCAGGCTTACAGCGCCCGCAGC 4571  
QY 609 -----LeuGlnTyrThrValGluValLysAs 617  
Db 4572 ACCATCAGCCAATAATACAGCTTCTCCGCTTCTTGCAGCTTACGCGGAGGTGAGGA 4631  
QY 617 nGlySerLeuIleIleProHisTyrAspHisTyrHisAsnIleLysPheGluTrpPheAs 637  
Db 4632 ACTGGCCACCTTC-----AATTCAAAAACAGAGGCGAGCGGA 4667  
QY 637 pGluGlyLeuTyrGluAlaProLysGlyTyrThrLeuGluAspLeuAlaThrValLy 657  
Db 4668 T-----GAGAAACCAAAG----- 4680  
QY 657 sTyrTyrValGluHisProAsnGluArgProHisSerAspAsnGlyPheGlyAsnAlaSe 677  
Db 4681 -----CTGGCGGAAGTCTCCAGTGCGACAGCCAAACCAA----- 4713  
QY 677 rAspHisValGlnArgAsnLysAsnGlyGlnAlaAspThrAsnGlnThrGluLysProSe 697  
Db 4714 -----ACCAAGAAAGCAGAGCAA-----CCAA 4739  
QY 697 rGluGluLysProGlnThrGluLysProGluGluGlu-----ThrProArgGluGluLy 715  
Db 4740 GCCAGAGCTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4799  
QY 715 sProGln-----SerGluLysProGluSerProLysPro-- 726  
Db 4800 CCCCCAGCTGGTGTCCCTTCCCTTGTGTCACAGCTCTCCCAAGCGCCCCCTCCACA 4859  
QY 727 -ThrGluGluProGluGluSerProGluGluSerGlu 738  
Db 4860 GTGCCCCCTTACCCAGCTCGAGCCCGCAGTCTCTCCAG 4896

Search completed: May 1, 2004, 08:41:51  
Job time : 1035 secs